Promoters: Prof. dr. ir. Pascal Boeckx
Laboratory of Applied Physical Chemistry (ISOFYS)
Faculty of Bioscience Engineering, Ghent University

Prof. dr. Bernard De Baets
Research Unit Knowledge-based Systems (KERMIT)
Faculty of Bioscience Engineering, Ghent University

Dean: Prof. dr. ir. Guido Van Huylenbroeck

Rector: Prof. dr. Paul Van Cauwenberge

ir. Dongmei Xue

Nitrate source classification in surface water via isotopic fingerprinting

Thesis submitted in fulfillment of the requirements
for the degree of Doctor (PhD) in Applied Biological Sciences

Dutch translation of the title:
Classificatie van nitraatbronnen in het oppervlaktewater via isotopische fingerprinting

To refer to this thesis:
Xue D. (2011) Nitrate source classification in surface water via isotopic fingerprinting

ISBN-number: 978-90-5989-437-2

Acknowledgements

23 February 2011

Pursuing a PhD project is a both painful and enjoyable experience. It is like climbing a high peak, step by step, accompanied by hardships, frustration, encouragement and trust and with a lot of help from kind-hearted people from ISOFYS and KERMIT. When I found myself at the top enjoying the beautiful view, I realized that it is the teamwork that got me there. Although it will not be enough to express my gratitude in words to all those people who helped me, I would still like to give my many, many thanks to all these people.

Particular thanks go to my promoters, Prof. Pascal Boeckx and Prof. Bernard De Baets, who accepted me as their PhD student. They offered me so much advice, patiently supervising me, and always guiding me in the right direction. Their wide knowledge and their logical ways of thinking have been of great value for me. Their understanding, encouraging and personal guidance have provided a good basis for the present thesis.

Special thanks are also given to Prof. Oswald van Cleemput, who provided encouragement and a lot of valuable comments on all my papers that have been published or in progress.

I appreciate the help from Jan, my IWT partner, who measured and corrected all of the isotope data I used in my thesis. Jorin, my IWT partner, also gave a lot of help for transporting, filtering, and analyzing samples. This thesis can not be fulfilled without their help.

I would also like to give my special thanks to Saskia, who helped me to deal with the financial reports for IWT, renew my contracts and residence card, and arrange my travel and accommodation for conferences. Warm thanks go to Katja, Eric, Dries, Samuel, Jeroen and Nasrin, who gave their individual supports and encouragements during my PhD.

I am very grateful for my parents. Their understanding and love encouraged me to work hard and to continue pursuing my PhD abroad. They helped me to take care of my baby that I can finish my thesis in time.

Last but not least, I dedicate this thesis to my husband (Dong Zhao) and my daughter (Xiaoge Zhao). They are the origin of my happiness. During my scientific research, my husband gave me incredible encouragement, support and patience, especially for cheering me up in difficult times. I owe my every achievement to my husband and my daughter.

Dongmei

Table of contents

List of symbols and abbreviations

N	nitrogen
NH_3	ammonia
NH_4^+	ammonium
NO_3^-	nitrate
NH_2OH	Hydroxylamine
NO_2^-	nitrite
N_2O	nitrous oxide
VMM	Flemish Environment Agency
MAP	Manure Action Plan
$^{15}N/^{14}N$	isotope ratio of nitrogen
$^{18}O/^{16}O$	isotope ratio of oxygen
$AgNO_3$	silver nitrate
δ	delta
‰	per mil
AIR	atmospheric air
VSMOW 2	Vienna Standard Mean Ocean Water 2
ε	enrichment factor
α	fractionation factor
class G	greenhouses in an agricultural area
class A	agriculture
class AGC	agriculture with groundwater compensation
class H	household
class AH	a combination of agriculture with horticulture
SIAR	stable isotope analysis in R
TC/EA-IRMS	thermal conversion/elemental analyzer-isotope ratio mass spectrometer
T	temperature
EC 20	electrical conductivity
DO	dissolved oxygen
O_{2sat}	oxygen saturation
Cl^-	chloride
PO_4^{3-}	phosphate

GC	gas chromatograph
TSB	tryptic soy broth
CF	correction factor
ICC	intraclass correlation coefficient
N_2O-AL	N_2O air liquid
KNO_3-IWS	KNO_3 internal working standard
Tukey HSD	Tukey Honest Significant Difference
NP	NO_3^- in precipitation
NF	NO_3^- fertilizer
NFR	NH_4^+ in fertilizer and rain
Soil	soil N
M&S	manure and sewage
W	winter
S	summer
CCI	percentage of correctly classified instances

List of tables

List of figures

Chapter 1:

Background, research objectives and description of study sites

1.1 The nitrogen cycle

The nitrogen (N) cycle is considered as one of the most important element cycles in ecosystems (Figure 1-1) (Nestler et al., 2011). Important processes in the N cycle include N_2 fixation, mineralization, nitrification, denitrification, anaerobic ammonium oxidation (anammox), dissimilatory nitrate reduction to ammonium (DNRA), volatilization, assimilation, etc.

The majority of N in ecosystems originally comes from the atmosphere, which stores N in a tremendous amount, mainly as N_2. Significant amounts of atmospheric N_2 enter soil by N-fixation processes which include lightning and natural bacterial fixation. The latter fixation process plays a predominant role for life. The bacteria processing the nitrogenase enzyme combine N_2 with hydrogen to produce ammonia (NH_3), which is then further converted by the bacteria to produce their own organic compounds. However, more and more atmospheric N_2 is now fixed by human activities, like energy production, fertilizer production and leguminous crop cultivation (Kendall, 1998).

In soil, N is primarily stored in organic matter. This organic N is converted to ammonium (NH_4^+) via a variety of bacteria, actinomycetes and fungi in a process called ammonification or mineralization.

The conversion of NH_4^+ to nitrate (NO_3^-) is known as autotrophic nitrification, which is the main oxidation pathway of NH_4^+. There are generally three steps in nitrification: firstly, NH_4^+ is oxidized to Hydroxylamine (NH_2OH) catalyzed by ammonia monooxygenase, in which O from O_2 is incorporated; secondly, NH_2OH is oxidized to nitrite (NO_2^-) catalyzed by hydroxylamine oxidoreductase, in which O from H_2O is incorporated; and finally, NO_2^- is oxidized to NO_3^- catalyzed by nitrite oxidoreductase with incorporation of O from H_2O, in which O can exchange between H_2O and NO_2^- or NO_3^- as a result of the reversibility of this step (Kool et al., 2007). In addition to autotrophic nitrification, organic N can also be oxidized to NO_3^- by heterotrophic bacteria, called heterotrophic nitrification.

Denitrification is a microbial process by which NO_3^- is converted to N_2 or nitrous oxide (N_2O) gas through a series of intermediate nitrogen oxide products. The process is performed primarily by heterotrophic bacteria (such as *Pseudomonas* and *Clostridium*) in anaerobic conditions and NO_3^- acts as a terminal electron acceptor for the oxidation of organic matter. However, chemo-autotrophic denitrification by bacteria (such as *Thiobacillus denitrificans*), which oxidizes sulfur, has also been identified (Batchelor and Lawrence, 1978).

Anaerobic ammonium oxidation (anammox) is an autotrophic process by which NH_4^+ is combined with NO_2^- under anaerobic conditions, producing N_2. The NO_2^- is derived from the partial autotrophic nitrification of NH_4^+ (nitritation).

Dissimilatory NO_3^- reduction to ammonium (DNRA) is a bacterial-mediated heterotrophic process occurring in anaerobic environments. There are two recognized DNRA pathways, fermentation and sulfur oxidation.

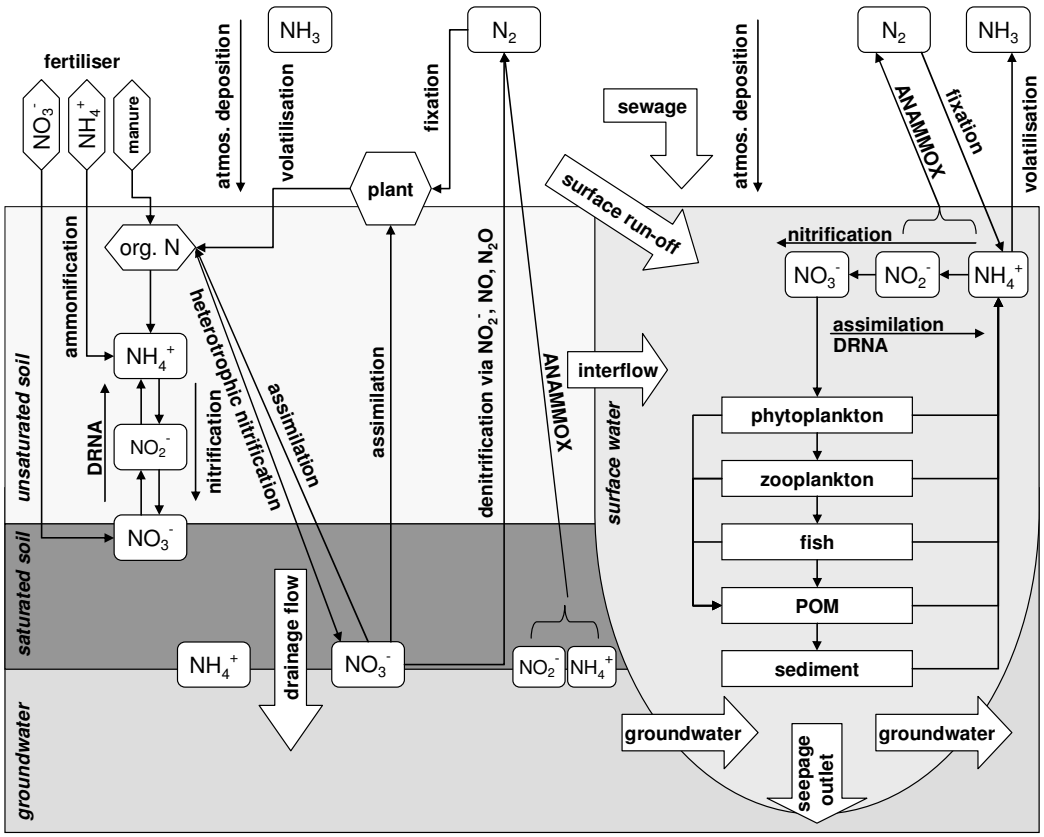


Figure 1-1: Nitrogen cycle (adapted from Nestler et al., 2011).

Plants can take up NH_4^+, NO_3^- and dissolved organic N from soil. NH_4^+ is used less by plants because it is extremely toxic at high concentrations. NO_3^- in plant is reduced to NO_2^- and then NH_4^+ for incorporation into amino acids, nucleic acids, and chlorophyll (Smil, 2000). This kind of process is regarded as assimilation.

1.2 Nitrate input in water via anthropogenic activities

Nitrogen is essential to all life processes as it forms amino acids, proteins, nucleic acids and DNA that are vital for all living cells. In natural systems, the main natural sources of N are bacteria that are able to fix N_2 from the air. Compared with the natural sources, anthropogenic activities have severely accelerated the N cycle, which results for instance in a global environmental concern for NO_3^- contamination in surface- and ground water. The World Health Organization (WHO) has set a limit NO_3^- concentration of 10 mg N L^{-1} for drinking water because high NO_3^- water increases risks of methemoglobinemia, commonly referred to as "blue baby syndrome" linked to brain damage and death by suffocation in infants (Spalding and Exner, 1993; Prasad and Power, 1995). In addition, high concentrations of NO_3^- in rivers and lakes can promote eutrophication that disrupts normal functioning of ecosystems, causing a variety of problems such as increased algae and aquatic plant growth, decreased water transparency, loss of desirable fish species because of dissolved oxygen depletion, etc. The major anthropogenic processes causing NO_3^- release include:

- **Overuse of N-based organic and inorganic fertilizers (Smil, 1999):** The application of N fertilizers to crops has caused increased leaching of NO_3^- into groundwater. A vast amount of N entering the groundwater system or direct runoff from the upper soil layer eventually flows into streams, rivers, lakes, and estuaries. In these systems, the added N is a dominant nutrient factor leading to eutrophication.
- **Elevated atmospheric N deposition (Benkovitz et al., 1996):** Fossil fuel (e.g. power plants and transportation) and biomass combustion increased emission of oxidized N (NO_x) and agriculture and intensive feedstock rearing increased emission of reduced N (NH_3) to the atmosphere. These forms can later be deposited and transported by storm water runoff.
- **Animal manure spreading:** Manure releases a large amount of gaseous ammonia (NH_3) to the atmosphere, and NH_4^+ which is derived from aqueous NH_3 can be subsequently oxidized to NO_3^-. This N enters the soil system and then the water system through leaching, groundwater flow, and runoff.
- **Discharge from septic tanks and sewage systems:** Sewage waste and septic tank can release large amounts of N by discharging through a drain field into the ground entering aquifers or directly discharging into surface water.

All of these anthropogenic activities increase the N load in ecosystems. Figure 1-2 demonstrates different source apportionments of N load in selected European countries and catchments (European Environmental Agency, 2005). These countries include: Austria, Belgium, Denmark, England/Wales, Estonia, Finland, Germany, Latvia, Lithuania, Netherlands, Norway, Poland and Sweden. The selected catchments contain: Axios, Danube, Daugava, Elbe, Ems, Odra, Po, Rhine, Vistula and Weser. The estimated sources include: point sources (such as discharges from urban wastewater, industry and fish farms), agriculture sources (such as fertilizer and manure application), background sources (natural land, for example forest) and total diffuse sources (agriculture sources plus background sources).

It is obvious that the total N load (size of the pie charts) is high in England/Wales, Netherlands, Belgium, Denmark and Germany. The Po river catchment in northern Italy also has a high load. The total N load is low in Eastern Europe, and even lower in the Nordic countries. For most selected European countries and catchments, agriculture or diffuse sources (agriculture plus background) account for more than 60% of the total load. However, for Norway and the Axios river, point sources contribute most of the anthropogenic load (the total N load is low). The background source of N is usually small compared with other sources. Thus, run-off from agricultural land is the principal source of N pollution in Europe.

Europe has been focusing on developing a *'full nitrogen approach'* that links the benefits, multiple environmental threats and policy issues. In particular, foundations are being developed through the NitroEurope Integrated project (the nitrogen cycle and its influence on the European greenhouse gas balance), the Nitrogen in Europe (NinE) programme of European Science Foundation, the COST Action 729 (Assessing and managing nitrogen fluxes in the atmosphere biosphere system in Europe), etc.

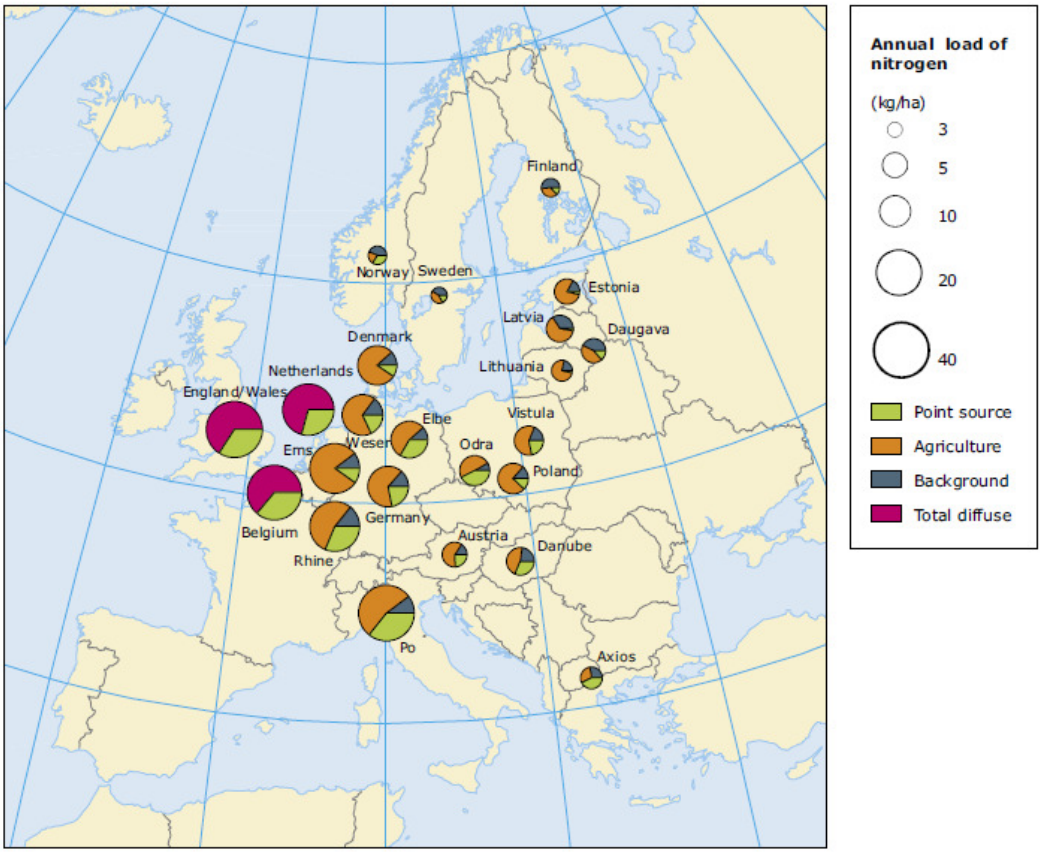


Figure 1-2: Source apportionment of N load in selected European countries and catchments (adapted from European Environmental Agency, 2005).

1.3 NO_3^- pollution in surface water in Flanders, Belgium

The implementation of the Nitrate Directive (EC, 2002) in Europe established a detailed framework for the protection of waters due to NO_3^- pollution from agricultural and imposed maximum allowable NO_3^- concentration of 50 mg NO_3^- L^{-1}. Thus, European countries design monitoring programmes based on the framework to identify water status, evaluate pressure on water systems, and detect water quality trends. The overall objective of the Nitrate Directive is the achievement of a 'good status' for all of Europe's surface- and ground waters within a 15-year period.

Flanders, which is situated in the northern part of Belgium, covers about 1,352,200 ha and the overall agricultural area is about 625,207 ha (Cazaux et al., 2007). The vast amount of N present in Flemish surface waters is assumed to originate from intensive manure and mineral fertilizer application in agriculture. VMM (Flemish Environment Agency) operates an extensive MAP (Manure Action Plan) monitoring network assessing the evolution of surface

water quality. This network is composed of about 800 sampling points on ditches, canals, brooks, rivers and lakes spread all over Flanders. Nitrate is monitored every month.

Figure 1-3 illustrates the NO_3^- pollution situation in surface water in Flanders in the period 2009 - 2010. The percentage shown in each basin is the percentage of monitoring sampling points in which NO_3^- concentrations were at least one time exceeding 50 mg NO_3^- L^{-1} in the period 2009–2010 in the corresponding basin. It is obvious that the pollution percentage number is different from basin to basin, from 0% in Dender basin to 69% in Yzer basin. The majority of pollution percentages is between 19% and 31%. The current pollution percentage for entire Flanders is 33%. Thus, NO_3^- pollution is still a serious problem in some basins like Leie (61%), Yzer (69%) and Meuse (47%).

Figure 1-4 demonstrates the NO_3^- pollution evolution in surface water in Flanders for the period 1999–2010. This figure shows the average pollution percentage and pollution variability (standard deviation) for all basins in Flanders. It is obvious that high pollution variability (20-25%) occurred during the whole monitoring period, which indicates that NO_3^- inputs vary largely in different basins. There is a clear decreasing trend on average between 1999 and 2003. This could result from the effectiveness of the MAP action with extensive monitoring and the limitation of fertilizer use. After 2003, the pollution percentage increased slightly and remained stable afterwards (around 30% to 40%).

Although monitoring NO_3^- concentration in water can provide useful information for water management, this approach requires a high density of monitoring sampling points in the long term and associated high cost. Furthermore, concentration data alone can not identify the specific NO_3^- sources (e.g. manure, fertilizer) and respective contribution of the sources responsible for water contamination. As a consequence, it is difficult to design and adjust environmental policies and management strategies to further reduce NO_3^- contamination in water. Therefore, we conclude that NO_3^- contamination in surface water in Flanders did not significantly reduce after 2003.

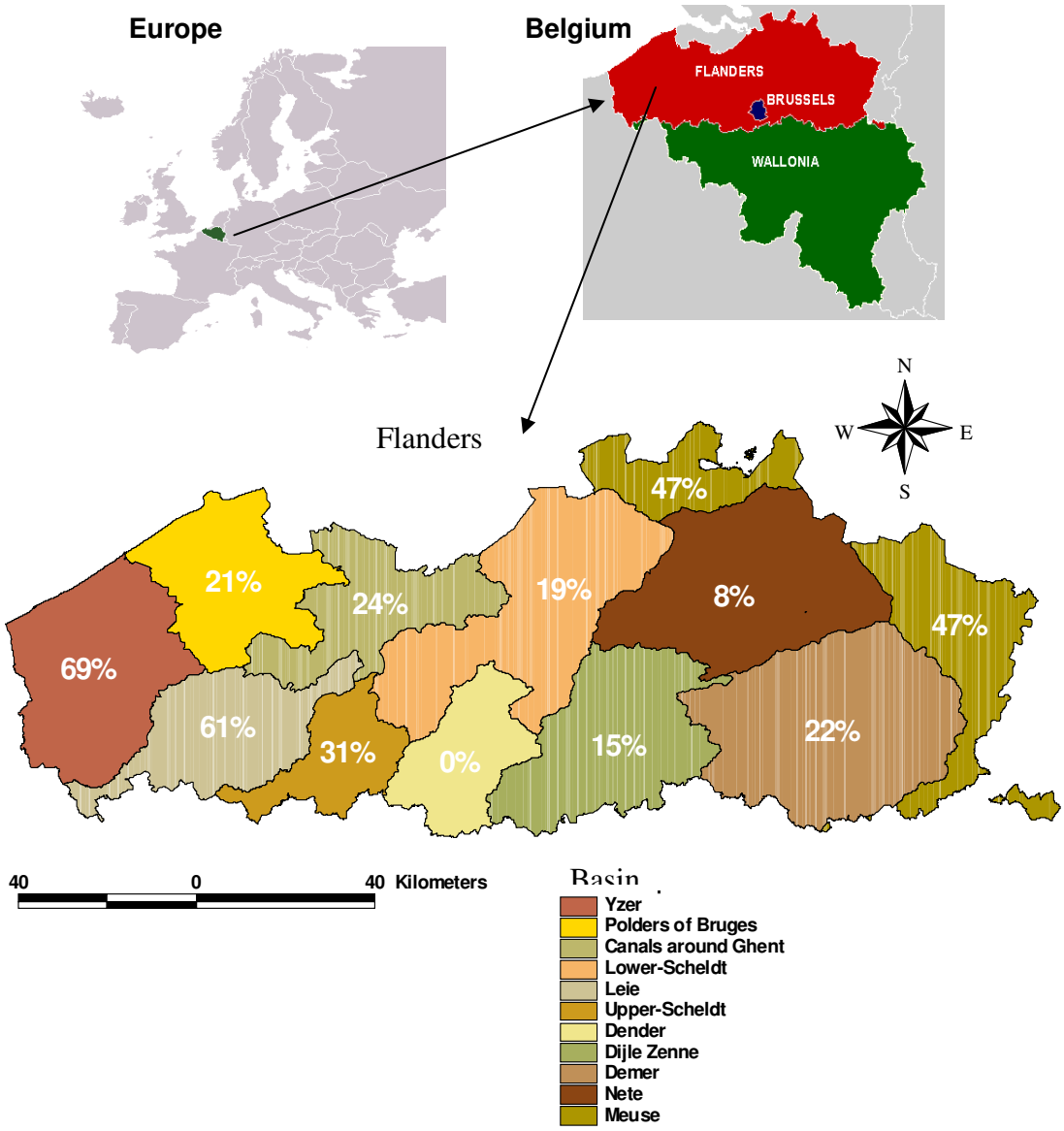


Figure 1-3: Pollution percentage of monitoring sampling points in each basin in Flanders in the period 2009-2010. Numbers indicate the percentage of monitoring sampling points in which NO_3^- concentrations were at least one time exceeding 50 mg NO_3^- L^{-1} in the period 2009-2010.

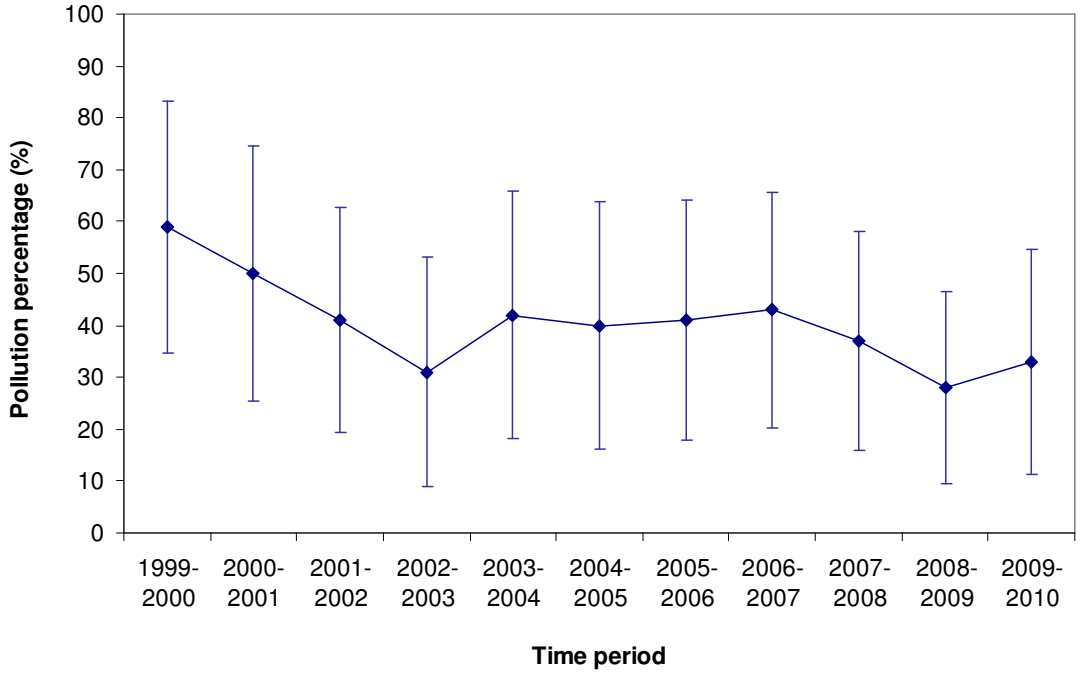


Figure 1-4: NO_3^- pollution percentage in surface water in Flanders in the period 1999–2010.

Recent research (Karr et al., 2001; Mayer et al. 2002; Spruill et al. 2002; Mitchell et al. 2003; Kaushal et al., 2006) has provided scientific proof that an isotope approach (using N and O isotopes of NO_3^-) is a useful tool to discriminate NO_3^- pollution sources, since different NO_3^- sources (fertilizer, manure, industrial/septic waste, atmospheric N deposition) have unique isotope ratios of N ($^{15}N/^{14}N$) and O ($^{18}O/^{16}O$). It is thus possible to identify these different sources via isotopic fingerprinting. In Chapter 2, a review regarding limitations and future prospects of N and O isotope application in water for NO_3^- source identification will be provided.

1.4 Research objectives and thesis outline

This thesis has the following objectives:

(1) To explore how to select representative sampling points for isotope monitoring based on a decision tree model performance;

(2) To compare the silver nitrate ($AgNO_3$) method and the bacterial denitrification method, both of which are frequently used analytical techniques to determine $\delta^{15}N$- (isotope ratio of N ($^{15}N/^{14}N$)) and $\delta^{18}O$-NO_3^- (isotope ratio of O ($^{18}O/^{16}O$)) in aqueous samples;

(3) To demonstrate how to correctly compute uncertainties on corrected $\delta^{15}N$- and $\delta^{18}O$-NO_3^- measurements in the bacterial denitrification method, as each step in the process potentially generates a relatively small amount of error depending on sample preparation, analytical conditions and equipment functioning;

(4) To identify and estimate different potential NO_3^- source contributions for 6 sampling points in surface water in Flanders based on 2-year $\delta^{15}N$- and $\delta^{18}O$-NO_3^- data analysis;

(5) To retrieve expert classification of 30 sampling points based on 2-year $\delta^{15}N$- and $\delta^{18}O$-NO_3^- and one year $\delta^{11}B$ data analysis.

Chapter 1 aims providing background information for NO_3^- contamination in surface and ground-water and describes the research objectives and study sites.

Chapter 2 aims providing background information for N and O stable isotope application, indicating both present limitation and future prospects for NO_3^-source identification.

Chapter 3 focuses on exploring a decision tree model based on physico-chemical data of 47 selected sampling points from the MAP monitoring work. These sampling points were selected and classified into different NO_3^- source classes (agriculture (class A), agriculture with groundwater compensation (class AGC), a combination of agriculture with horticulture (class AH), greenhouses in an agricultural area (class G) and households (class H)) by expert knowledge from VMM. Thirty representative sampling points (six sampling points per class) were selected based on the decision tree model performance for further isotope monitoring and retrieve of expert classification.

Chapter 4 describes two frequently used analytical techniques, the silver nitrate ($AgNO_3$) method and the bacterial denitrification method to measure $\delta^{15}N$- and $\delta^{18}O$-NO_3^-. A thorough method comparison was carried out using a relatively large number of real surface water samples (42 water samples) having a wide range of $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values.

Chapter 5 demonstrates a correct way to compute uncertainties on corrected $\delta^{15}N$- and $\delta^{18}O$-NO_3^- via the bacterial denitrification method. Since errors can be potentially produced by preparing samples, changing analytical conditions and running equipment, it is necessary to consider the uncertainty on the final result after correction;

Chapter 6 focuses on identifying and evaluating NO_3^- source inputs for 30 sampling points based on 2-year isotope monitoring data. Qualitative information regarding predominant NO_3^- sources was obtained by a bi-plot approach ($\delta^{15}N$-NO_3^- vs. $\delta^{18}O$-NO_3^-), but further quantitative information has been estimated by a "fingerprint" tool, SIAR (stable isotope analysis in R) (Parnell et al., 2010).

Chapter 7 conducts retrieving expert classification of 30 sampling points using k-means clustering method based on $\delta^{11}B$ data and SIAR outputs for all 30 sampling points.
Finally, the main findings and conclusions from this study and future research are summarized in Chapter 8.

1.5 Description of study sites

The study sites are located in Flanders, the northern part of Belgium. The prevalent climate is temperate maritime, influenced by the North Sea and Atlantic Ocean, with cool summers and moderate winters. Flanders covers about 1,352,200 ha and the overall agricultural area is about 625,207 ha (Cazaux et al., 2007). The agriculture area is covered with meadows, pasture, fodder crops and arable farming, which indicates the importance of livestock in Flanders (Smets et al., 2004). The main crops are cereals, potatoes and sugar beet. In addition, horticulture activities occupy 8% of the agricultural area, half for growing vegetables (Smets et al., 2004).
The on-going nitrate management programme, the MAP, uses about 800 surface water sampling points for monitoring water quality evolution. Experts from VMM (Flemish Environmental Agency) are responsible for managing those sampling points. Thus, 3 meetings were organized between VMM and Gent University research team to select representative 47 sampling points from the MAP according to historical concentration variability, land use types and field knowledge. These sampling points are distributed in different basins over the whole of Flanders (Figure 1-5). Experts from VMM classified the sampling points into 5 different NO_3^- source classes based on different land use types: 7 sampling points were classified into agriculture (class A) class; 15 sampling points were classified into agriculture with groundwater compensation (class AGC) class; 6 sampling points were classified into a combination of agriculture with horticulture (class AH) class; 11 sampling points were classified into greenhouses in an agricultural area (class G) class and 8 sampling points were classified into household (class H) class. The surface water sampling was done by VMM as an additional sampling during monthly monitoring of these locations. Surface water was taken as a grab sample, transported to the nearest sampling station center and stored frozen in 1L HDPE (High Density Polyethylene) bottles until the water samples were finally transported to the ISOFYS laboratory of Gent University for determination of physico-chemical properties and $\delta^{15}N$- and $\delta^{18}O$-NO_3^-.

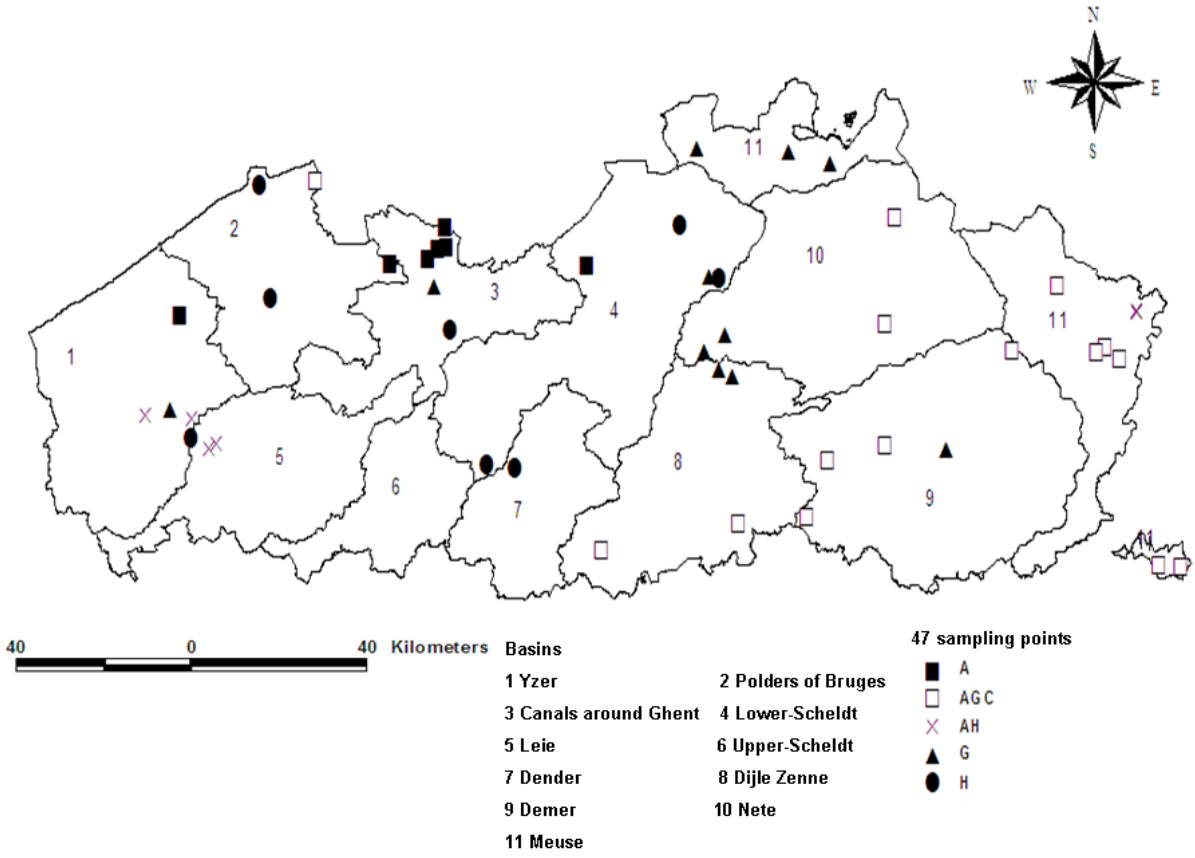
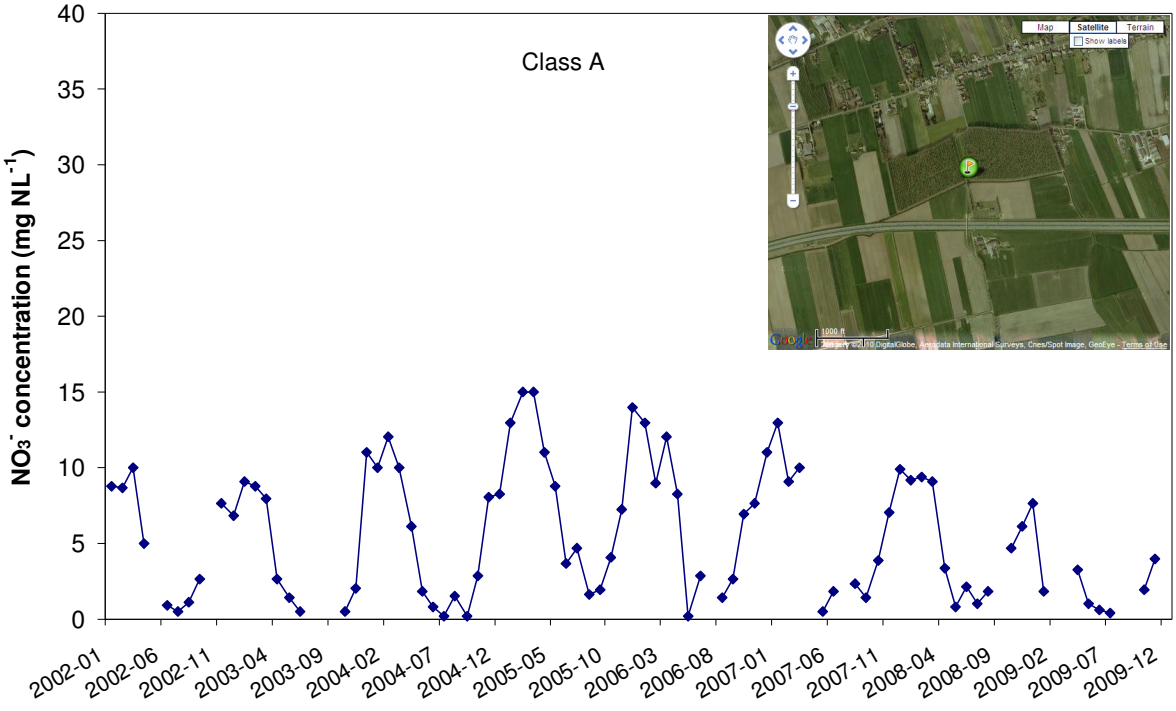
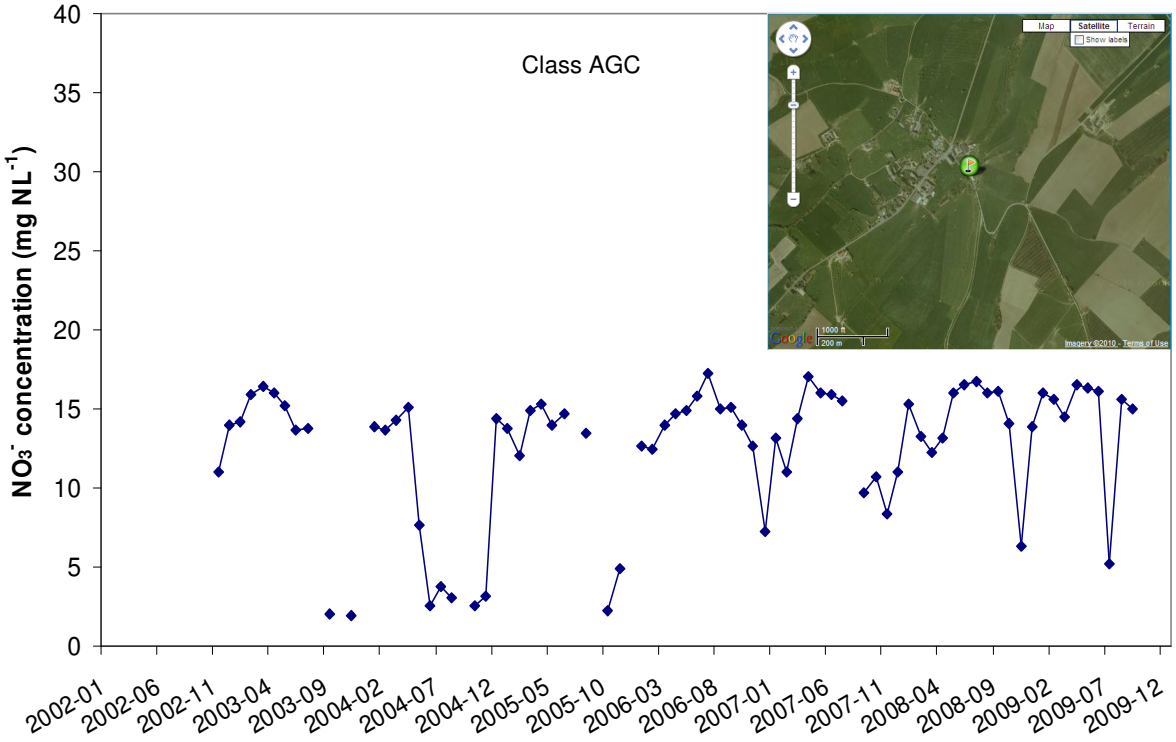
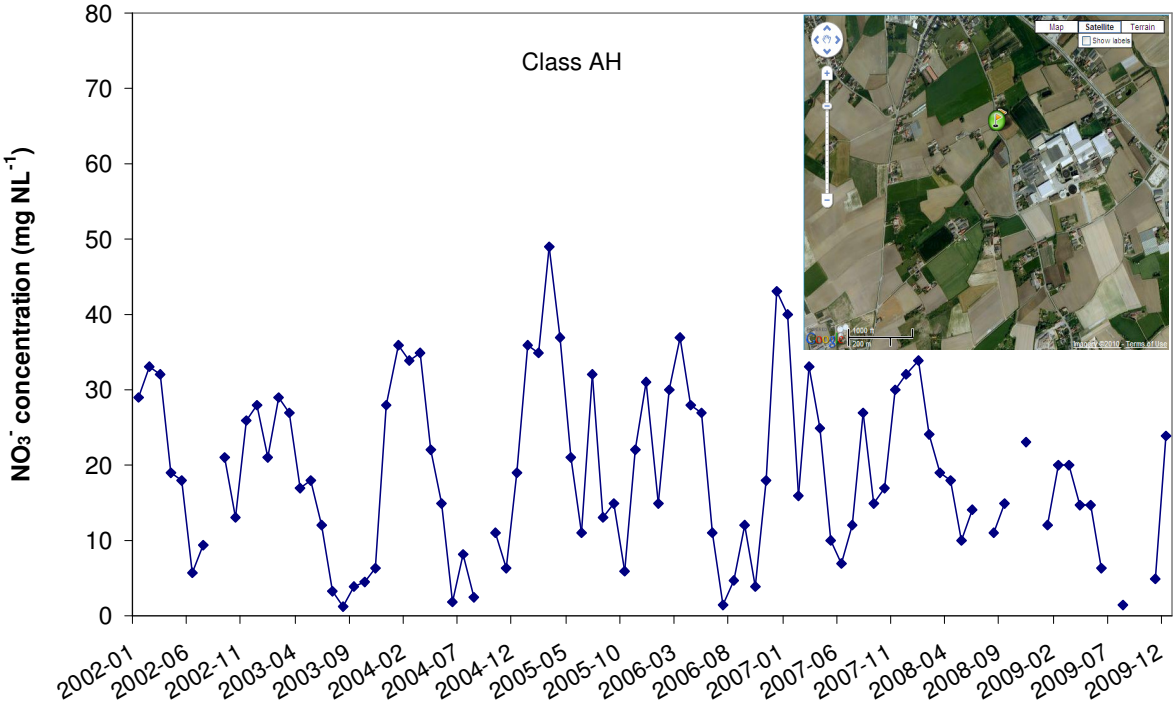


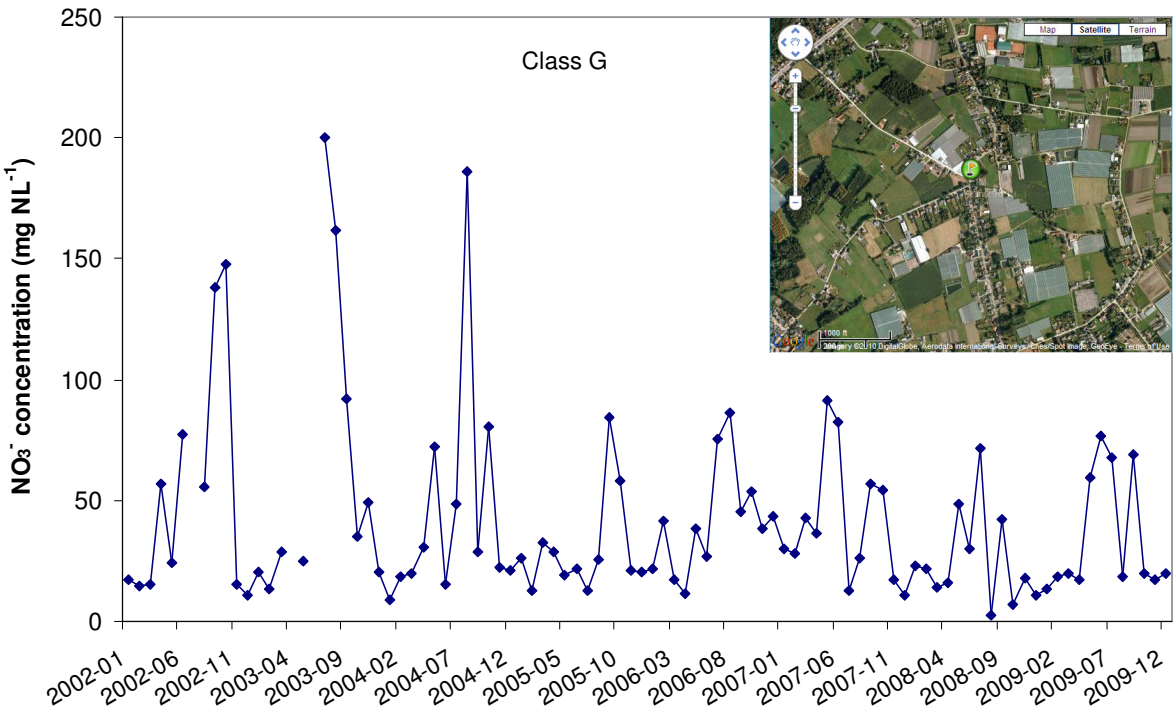
Figure 1-5: Distribution and classification of 47 sampling points in Flanders. A represents agriculture; AGH represents agriculture with groundwater compensation; AH represents a combination of agriculture with horticulture; G represents greenhouses in an agricultural area; H represents households.

Furthermore, NO_3^- concentration evolution behaviors is different for different land use types. Figure 1-6 illustrates the NO_3^- concentration evolution (from 2002 to 2009) for five sampling points in 5 different classes.


Class A
NO3- concentration (mg NL-1)
0
5
10
15
20
25
30
35
40
2002-01
2002-06
2002-11
2003-04
2003-09
2004-02
2004-07
2004-12
2005-05
2005-10
2006-03
2006-08
2007-01
2007-06
2007-11
2008-04
2008-09
2009-02
2009-07
2009-12


Class AGC
NO3- concentration (mg NL-1)
0
5
10
15
20
25
30
35
40
2002-01
2002-06
2002-11
2003-04
2003-09
2004-02
2004-07
2004-12
2005-05
2005-10
2006-03
2006-08
2007-01
2007-06
2007-11
2008-04
2008-09
2009-02
2009-07
2009-12


Class AH
NO3- concentration (mg NL-1)
0
10
20
30
40
50
60
70
80
2002-01
2002-06
2002-11
2003-04
2003-09
2004-02
2004-07
2004-12
2005-05
2005-10
2006-03
2006-08
2007-01
2007-06
2007-11
2008-04
2008-09
2009-02
2009-07
2009-12


Class G
NO3- concentration (mg NL-1)
0
50
100
150
200
250
2002-01
2002-06
2002-11
2003-04
2003-09
2004-02
2004-07
2004-12
2005-05
2005-10
2006-03
2006-08
2007-01
2007-06
2007-11
2008-04
2008-09
2009-02
2009-07
2009-12

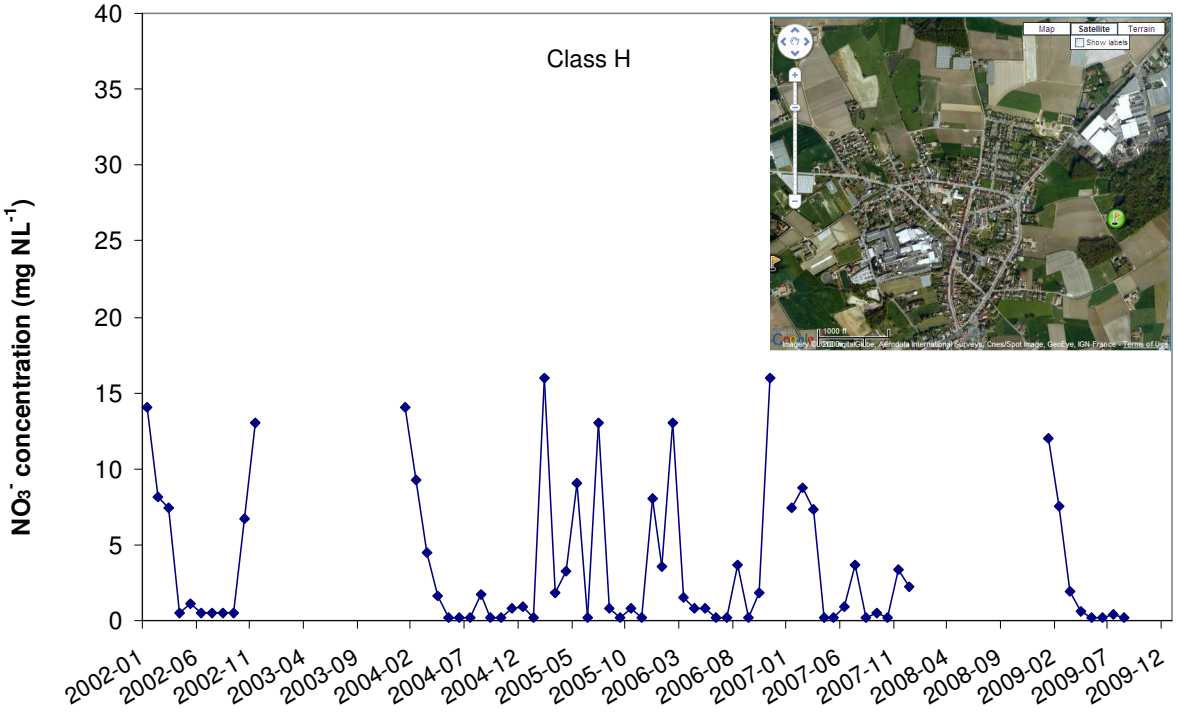


Figure 1-6: A view of location and NO_3^- concentration evolution (from 2002 to 2009) of 5 sampling points in 5 different land use areas.

For class A, the selected sampling point is located in an agricultural area. The NO_3^- concentrations ranged between 0.2 to 15 mg N L^{-1}. The NO_3^- concentration evolution reveals a periodic pattern that yearly highest concentration occurred in winter time. A possible reason is that diminished plant uptake and rapid transportation of water through the soil during wet and mild winter (climatic characteristics of Belgium) increased the NO_3^- inputs into surface water.

The selected sampling point in class AGC is also located in an agricultural area, but with ground water compensation. No apparent seasonal variation has been observed and the NO_3^- concentration varied in a stable trend. The majority of the NO_3^- concentrations were in a range from 6.3 to 17.3 mg N L^{-1} except for several low ones.

For class AH, the selected sampling point is located in an agricultural area with horticultural activity. The NO_3^- concentrations varied from 1.2 to 49.0 mg N L^{-1} and multiple concentration peaks have been found. The yearly highest concentration occurred in winter time.

The selected sampling point in class G is located in the vicinity of greenhouses in an agricultural area. The NO_3^- concentrations varied widely, from 2.4 to 200 mg N L^{-1}. The extremely high NO_3^- concentrations mainly occurred in summer time, which could be a result

from intensive greenhouse activities in summer time and a large amount of surplus mineral fertilizers used in greenhouses.

The sampling point in class H is located in a household area. The NO_3^- concentrations varied between 0.2 and 16 mg N L^{-1} without obvious seasonal trend. There were two breaks of two years during NO_3^- concentration monitoring, which were determined by VMM as “resting point” without sampling.

Chapter 2:

Present limitations and future prospects of stable isotope methods for nitrate source identification in surface- and groundwater

This chapter has been edited from:

Xue, D., Botte, J., De Baets, B., Accoe, F., Nestler, A., Taylor, P., Van Cleemput, O., Berglund, M. and Boeckx, P. (2009) Present limitations and future prospects of stable isotope methods for nitrate source identification in surface- and groundwater. Water Res. 43, 1159–1170.

2.1 Abstract

Nitrate (NO_3^-) contamination of surface- and groundwater is an environmental problem in many regions of the world with intensive agriculture and high population densities. Knowledge of the sources of NO_3^- contamination in water is important for better management of water quality. Stable nitrogen ($\delta^{15}N$) and oxygen ($\delta^{18}O$) isotope data of NO_3^- have been frequently used to identify NO_3^- sources in water. This chapter summarizes typical $\delta^{15}N$- and $\delta^{18}O$-NO_3^- ranges of known NO_3^- sources, interprets constraints and future outlooks to quantify NO_3^- sources, and describes two frequently used analytical techniques ("ion-exchange method" and "bacterial denitrification method") for $\delta^{15}N$- and $\delta^{18}O$-NO_3^- determination. Isotope data can provide evidence for the presence of dominant NO_3^- sources. However, quantification, including uncertainty assessment, is lacking when multiple NO_3^- sources are present. Moreover, fractionation processes are often ignored, but may largely constrain the accuracy of NO_3^- source identification. These problems can be overcome if (1) NO_3^- isotope data are combined with co-migrating discriminators of NO_3^- sources (e.g. ^{11}B), which are not affected by transformation processes, (2) contributions of different NO_3^- sources can be quantified via linear mixing models (e.g. SIAR), and (3) precise, accurate and high throughput isotope analytical techniques become available.

2.2 Introduction

Nitrate (NO_3^-) contamination in water is an environmental problem worldwide. Increased input of reactive nitrogen (N) is attributed to intensive land use, increased use of N-containing organic and inorganic fertilizers (Smil, 1999), animal manure, discharge of human sewage and elevated atmospheric N deposition (Benkovitz et al., 1996). The World Health Organization has set a limit NO_3^- concentration of 10 mg N L^{-1} for drinking water. The implementation of the Nitrate Directive (EC, 2002) in Europe established a detailed framework for prevention of NO_3^- pollution to waters. However, to effectively control NO_3^- contamination in water, NO_3^- source inputs should be better understood. Subsequently, dedicated measures could be applied to prevent or minimize contamination.

Stable N isotope data of NO_3^- ($\delta^{15}N$-NO_3^-) have been frequently used to estimate the origin of NO_3^- in water (Wells and Krothe, 1989; Feast et al., 1998; Mayer et al., 2002), because the isotope composition of N in NO_3^- is generally different among various NO_3^- sources such as atmospheric N_2, soil, chemical fertilizers, and manure. However, the origin of NO_3^- must be

linked to the entire N cycle, since values of $\delta^{15}N\text{-}NO_3^-$ can be biased due to mixing of distinct NO_3^- sources and kinetic isotopic fractionation (e.g. denitrification) (Kellman and Hillaire-Marcel, 1998). Hence, the $\delta^{15}N$ signature alone does not allow for conclusive identification of NO_3^- sources. Therefore, a dual isotope approach (combination of ^{15}N and ^{18}O) has been widely used and provides more conclusive information for tracing sources of NO_3^- in water (Komor, 1997; Aravena and Robertson, 1998; Widory et al., 2004; Seiler, 2005). In addition, the analytical methodologies for both $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ have been improved considerably in recent years and are ready to become high throughput analytical techniques.

This chapter is composed of four parts: (1) basic information on N and O isotopes and isotopic ranges of known NO_3^- sources, (2) factors affecting the isotopic composition of NO_3^- sources, (3) outlooks to improve the identification of NO_3^- sources, and (4) description and evaluation of analytical techniques for $\delta^{15}N$ and $\delta^{18}O$ determination in NO_3^-. Thus, the overall objective is to provide a comprehensive overview of the current state-of-the-art to identify NO_3^- sources in water via $\delta^{15}N$ and $\delta^{18}O\text{-}NO_3^-$ data and to provide an outlook for future improvements for NO_3^- source identification.

2.3 Stable nitrogen and oxygen isotopes and kinetic isotope fractionation

There are two naturally occurring stable isotopes of nitrogen (N), ^{14}N and ^{15}N. The majority of N in the atmosphere is composed of ^{14}N (99.6337%) and the remainder is composed of ^{15}N (0.3663%) (Junk and Svec, 1958). Oxygen (O) is composed of three stable isotopes, ^{16}O (99.759%), ^{17}O (0.037%), and ^{18}O (0.204%) (Cook and Lauer, 1968). Stable isotope ratios are usually expressed in delta (δ) units and a per mil (‰) notation relative to the respective international standards:

$$\delta_{\text{sample}}\,(‰) = \left(\frac{R_{sample}}{R_{standard}} - 1\right) \times 1000 \qquad (2\text{-}1)$$

where R is the $^{15}N/^{14}N$ or $^{18}O/^{16}O$ ratio of the sample and standard, respectively for $\delta^{15}N$ and $\delta^{18}O$. $\delta^{15}N$ values are reported relative to atmospheric air (AIR) and $\delta^{18}O$ values are reported relative to Vienna Standard Mean Ocean Water 2 (VSMOW 2). When δ_{sample} is positive, it indicates enrichment in the heavy isotope. A negative δ_{sample} indicates depletion in the heavy isotope.

Stable isotopes are subjected to kinetic isotope fractionation, which results in depleted instantaneous products, while the remaining substrate becomes gradually more enriched with the heavy isotope. The latter is characterized by an enrichment factor (ε) in ‰ which is defined as:

$$\varepsilon = 10^3 (\alpha - 1) \quad (2\text{-}2)$$

in which α is the fractionation factor and is defined as

$$\alpha = \kappa_{^{15}N} / \kappa_{^{14}N} \quad \text{or}$$
$$\alpha = \kappa_{^{18}O} / \kappa_{^{16}O} \quad (2\text{-}3)$$

where κ are the rate constants in the transformation process. In a closed system, isotopic enrichment can be expressed by the following Rayleigh equation:

$$\delta_{S(t)} = \delta_{S0} + \varepsilon \ln (S_t/S_0) \quad (2\text{-}4)$$

where ε is the isotopic enrichment factor, δ_{S0} and $\delta_{S(t)}$ represent the isotopic compositions of the substrate at time 0 and t, and S_0 and S_t are the concentrations of the substrate at time 0 and t, respectively. This equation can be used to estimate the enrichment factor ε by plotting measured values of $\delta_{S(t)}$ against $\ln (S_t/S_0)$.

2.4 $\delta^{15}N$ values of NO_3^- sources

The use of $\delta^{15}N\text{-}NO_3^-$ for identification of NO_3^- sources dates back to the 1970s. The first study using $\delta^{15}N\text{-}NO_3^-$ was reported by Kohl et al. (1971) for estimation of fertilizer contribution to NO_3^- in the Sangamon River (Illinois, USA). Different NO_3^- sources can be discriminated from each other as NO_3^- originating from different sources shows characteristic $\delta^{15}N$ values. Boxplots of $\delta^{15}N$ data for various N sources and sinks are displayed in Figure 2-1, which shows percentile values and outliers of $\delta^{15}N$ for each NO_3^- source. The typical $\delta^{15}N$ ranges are presented by the 10th and 90th percentiles.

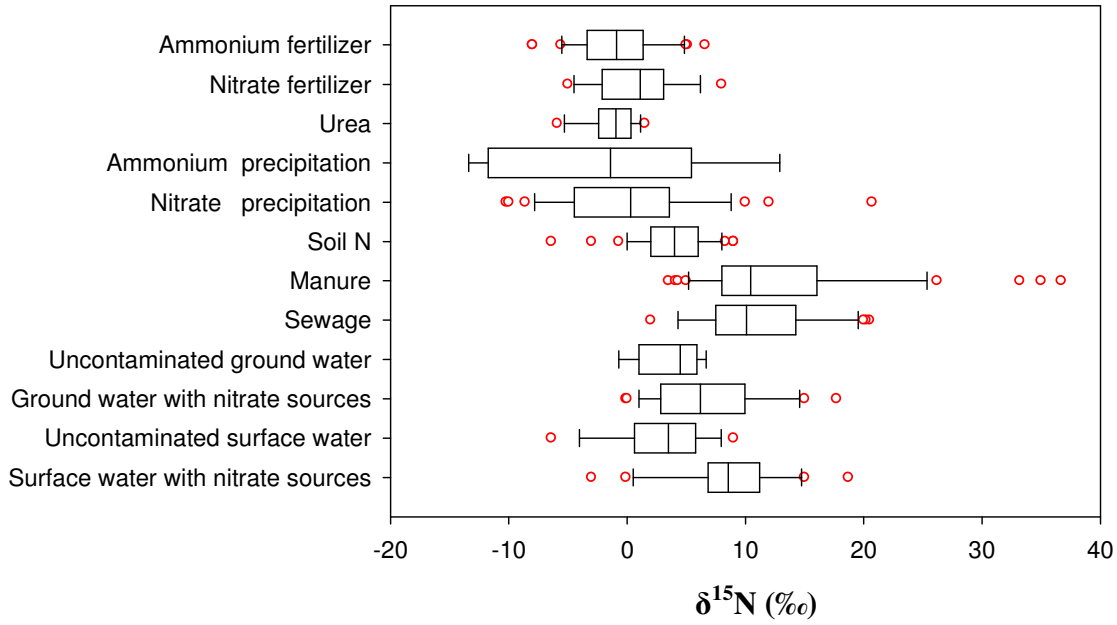


Figure 2-1: Boxplots of $\delta^{15}N$ values of NO_3^- from various sources and sinks. Boxplots illustrate the 25th, 50th and 75th percentiles; the whiskers indicate the 10th and 90th percentiles; and the circles represent outliers. NH_4^+ fertilizer- Heaton (1986); Hübner (1986); Cao et al.(1991); Wassenaar (1995); Kendall (1998); Roadcap et al. (2001); Choi et al. (2003b); Singleton et al. (2007); Bateman and Kelly (2007); Choi et al. (2007); Li et al. (2007); NO_3^- fertilizer- Black and Waring (1977); Hübner (1986); Hirata (1996); Kendall (1998); Roadcap et al. (2001); Deutsch et al. (2006a); Bateman and Kelly (2007); Singleton et al. (2007); Urea- Black and Waring (1977); Heaton (1986); Cao et al. (1991); Wassenaar (1995); Choi et al. (2002a); Deutsch et al. (2006a); Bateman and Kelly (2007); Choi et al. (2007); Li et al. (2007); NH_4^+ precipitation- Hübner (1986); Hirata (1996); Russell et al. (1998); Li et al. (2007); Zhang et al. (2008); NO_3^- precipitation- Kendall (1998); Russell et al. (1998); Spoelstra et al. (2001); Williard et al. (2001); Pardo et al. (2004); Kellman (2005); Piatek et al. (2005); Deutsch et al. (2006a); Finlay et al. (2007); Hales et al. (2007); Li et al. (2007); Singleton et al. (2007); Townsend-Small et al. (2007); Lee et al. (2008); Zhang et al. (2008); Soil N- Heaton (1986); Mariotti et al. (1988); Wassenaar (1995); Lindau et al. (1997); Fogg et al. (1998); Kendall (1998); McClelland and Valiela (1998); Mayer et al. (2001); Williard et al. (2001); Kellman (2005); Spoelstra et al. (2007); Singleton et al. (2007); Manure- Wassenaar (1995); Fogg et al. (1998); Kendall (1998); Karr et al. (2001); Choi et al. (2002b); Choi et al. (2003a); Choi et al. (2003b); Curt et al. (2004); Widory et al. (2004); Kellman (2005); Widory et al. (2005); Choi et al. (2007); Sewage- Hirata (1996); Fogg et al. (1998); Kendall (1998); Mariotti et al. (1988); McClelland and Valiela (1998); Curt et al.

(2004); Widory et al. (2004); Widory et al. (2005); Li et al. (2007); Uncontaminated groundwater- Burg and Heaton (1998); Brandes et al. (1998); Rock and Mayer (2002); Seiler (2005); Deutsch et al. (2006a); Choi et al. (2007); Groundwater with nitrate sources- Iqbal et al. (1997); Fogg et al. (1998); Mengis et al. (1999); Panno et al. (2001); Min et al. (2002); Choi et al. (2002a); Mitchell et al. (2003); Katz et al. (2004); Jun et al. (2005); Li et al. (2007); Choi et al. (2007); Uncontaminated surface water- Spoelstra et al. (2001); Williard et al. (2001); Mayer et al. (2002); Pardo et al. (2004); Piatek et al. (2005); Deutsch et al. (2006b); Hales et al. (2007); Lee et al. (2008); Surface water with nitrate sources- Rock and Mayer (2002); Pardo et al. (2004); Kellman (2005); Deutsch et al. (2006a); Deutsch et al. (2006b); Voss et al. (2006); Lee et al. (2008).

Ammonium (NH_4^+) fertilizer, NO_3^- fertilizer and urea are (in)organic fertilizers, which are produced by fixation of atmospheric N_2 and show small differences in $\delta^{15}N$ content as a result of small fractionation during subsequent processing of the fixed N (Flipse and Bonner, 1985). These (in)organic fertilizers show a typical $\delta^{15}N$ value range (10th and 90th percentiles in boxplot) between -6‰ and +6‰.

The typical $\delta^{15}N$ values for atmospheric N deposition are situated between -13‰ and +13‰. This is controlled by complex chemical reactions in the atmosphere and various anthropogenic sources such as combustion of fossil fuels (Hübner, 1986; Kendall, 1998).

Manure and sewage are enriched in ^{15}N relative to other N sources. During storage, treatment and application of sewage and animal wastes, ammonia (NH_3) volatilization causes a large enrichment of ^{15}N in the residual NH_4^+. This NH_4^+ is subsequently converted into ^{15}N-enriched NO_3^-. Thus, $\delta^{15}N$ values of NO_3^- originating from manure are between +5‰ and +25‰ and sewage between +4‰ and +19‰.

The typical $\delta^{15}N$ values of soil N range from 0‰ to +8‰. This is linked to the relative rate of mineralization and nitrification. Other factors including soil depth, vegetation, climate and site history may also affect the $\delta^{15}N$ values of soil (Kendall, 1998; Mayer et al., 2001).

Nitrogen contamination in surface- and groundwater (recognized as NO_3^- sinks) in agricultural regions is generally attributed to the application of fertilizers, animal waste and other organic residues. Additionally, intensive land use, vegetation and industrial activity are also important contributors of NO_3^- to surface- and groundwater. Therefore, the $\delta^{15}N$ values of NO_3^- in surface- and groundwater typically vary from -4‰ to +15‰.

2.5 $\delta^{18}O$ values of NO_3^- sources

Values of $\delta^{18}O$-NO_3^- have increasingly been used as an additional means to identify sources of NO_3^- in water (Kendall et al., 1996; Mayer et al., 2002; Pardo et al., 2004; Deutsch et al., 2006a). Durka et al. (1994) found that $\delta^{18}O$ was more useful than $\delta^{15}N$ to separate atmospheric NO_3^- deposition from microbially produced soil NO_3^-, because $\delta^{18}O$ signatures of atmospheric NO_3^- (from +52.5‰ to +60.9‰) and microbially-produced soil NO_3^- (from +0.8‰ to +5.8‰) differ significantly. Wassenaar (1995) also suggested that $\delta^{18}O$ may be a good tracer of NO_3^- sources, because it distinguishes synthetic NO_3^- fertilizers from other NO_3^- sources.

Boxplots of $\delta^{18}O$ signatures of NO_3^- sources are displayed in Figure 2-2. It is clear that typical $\delta^{18}O$ values of NO_3^- from nitrification (including $\delta^{18}O$ values from microbial production of NH_4^+ in fertilizer and precipitation, NO_3^- derived from soil N and NO_3^- derived from manure and sewage) are lower than that of NO_3^- from precipitation and NO_3^- in fertilizer.

The $\delta^{18}O$ of microbially produced NO_3^-, NO_3^- in precipitation and NO_3^- fertilizer are regulated by different processes. The $\delta^{18}O$ of microbially produced NO_3^- is determined by the $\delta^{18}O$ of H_2O and atmospheric O_2. In theory, two oxygen atoms of the newly formed NO_3^- are derived from H_2O oxygen, with the third oxygen atom incorporated from atmospheric O_2 (Hollocher, 1984; Anderson and Levine, 1986; Böhlke et al., 1997; Durka et al., 1994; Kendall, 1998; Wassenaar, 1995; Mayer et al., 2001). NO_3^- derived from nitrification should theoretically have $\delta^{18}O$ values between -10‰ and +10‰, since environmental water typically has $\delta^{18}O$ values between -25‰ and +4‰ and $\delta^{18}O$ of atmospheric O_2 is +23.5‰ (Hollocher, 1984; Durka et al., 1994; Kendall, 1998). It has, however, been observed that $\delta^{18}O$ values of microbially produced NO_3^- are up to 5‰ higher than the calculated theoretical maximum of 10‰ (Aravena et al., 1993; Kendall, 1998; Mayer et al., 2001). This can be expected if (1) the $\delta^{18}O$ value of H_2O is isotopically enriched as a result of evaporation, (2) the $\delta^{18}O$ value of soil O_2 is higher than that of atmospheric O_2 resulting from O isotope fractionation during respiration, (3) there is significant isotope fractionation during the incorporation of oxygen from H_2O and O_2 into the newly formed NO_3^-, (4) the ratio of oxygen incorporation from H_2O and O_2 is not 2:1 (e.g. more O_2 may be derived from atmospheric O_2 when NH_4^+ is limiting), or (5) a different bacterial process dominates the nitrification reaction and utilizes a greater amount of atmospheric O_2 in low pH environments (Aravena et al., 1993; Wassenaar, 1995; Böhlke et al., 1997; Kendall, 1998; Mayer et al. 2001). Hence, $\delta^{18}O$-NO_3^- produced is relatively enriched compared to the theoretical calculation for nitrification.

The $\delta^{18}O\text{-}NO_3^-$ of precipitation is controlled by complex atmospheric processes, which result in large spatial and temporal variability in $\delta^{18}O\text{-}NO_3^-$. Affecting factors include isotopic fractionation during NO_3^- formation caused by thunderstorms, the isotopic signature of the reactive oxygen in the atmosphere that combines with NO_x to form NO_3^-, and any isotopic fractionation during reactions in the atmosphere (Kendall, 1998; Pardo et al., 2004). Thus, $\delta^{18}O$ of atmospheric NO_3^- is enriched relative to atmospheric O_2 (+23.5‰), showing a wide range between +25‰ and +75‰.

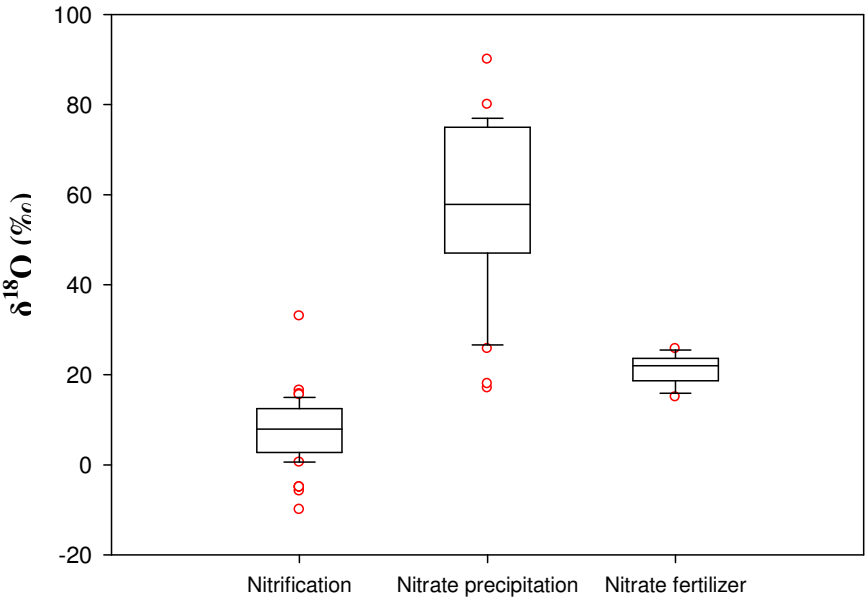


Figure 2-2: Boxplots of $\delta^{18}O$ values of NO_3^- generated during nitrification, nitrate precipitation and nitrate fertilizer. Box plots illustrate the 25th, 50th and 75th percentile; the whiskers indicate the 10th and 90th percentiles; and the circles represent data outliers. Nitrification- Hübner (1986); Wassenaar (1995); Kendall (1998); Mayer et al. (2001); Spoelstra et al. (2001); Williard et al. (2001); Mayer et al. (2002); Rock and Mayer (2002); Fukada et al. (2003); Pardo et al. (2004); Piatek et al. (2005); Seiler (2005); Deutsch et al. (2006a); Hales et al. (2007); Singleton et al. (2007); Spoelstra et al. (2007); Lee et al. (2008); Umezawa et al. (2008); Nitrate precipitation- Kendall (1998) ; Spoelstra et al. (2001); Williard et al. (2001); Pardo et al. (2004); Piatek et al. (2005); Deutsch et al. (2006a); Finlay et al. (2007); Hales et al. (2007); Singleton et al. (2007); Lee et al. (2008); Nitrate fertilizer- Hübner (1986); Kendall (1998); Roadcap et al. (2001); Deutsch et al. (2006a); Singleton et al. (2007).

Synthetic NO_3^- fertilizers have distinctive $\delta^{18}O$ values (Amberger and Schmidt, 1987), with a range from +17‰ to +25‰, as the O is mainly derived from atmospheric O_2 ($\delta^{18}O$ = 23.5‰).

2.6 Effect of $\delta^{18}O$ on NO_3^- source identification

Since $\delta^{18}O$-NO_3^- is used as an additional marker of NO_3^- sources, it is meaningful to assess the effect size of $\delta^{18}O$-NO_3^- data on $\delta^{15}N$-NO_3^- for NO_3^- source identification. An effect size is a measure of the magnitude of the relationship between two variables. The calculation of the effect size is carried out by the *d*-index of Hedges implemented in the Meta Win software (Rosenberg et al., 2000). It is calculated as:

$$d = \frac{\overline{X}_e - \overline{X}_c}{S} J \tag{2-5}$$

where $\overline{X}_e$ is the mean of the experimental group ($\delta^{18}O$-NO_3^-), $\overline{X}_c$ is the mean of the control group ($\delta^{15}N$-NO_3^-), *S* is the pooled standard deviation, and *J* is a factor that corrects for small sample bias. An effect size value (Hedges'*d*) equal to zero indicates no effect, and effect sizes of 0.2, 0.5 and 0.8 or greater indicate small, medium and large effects, respectively. Values below zero indicate negative effects.

The effect size of $\delta^{18}O$-NO_3^- data on $\delta^{15}N$-NO_3^- for NO_3^- source identification is shown in Figure 2-3. It is clear that some effect size values are above zero, indicating a positive contribution of $\delta^{18}O$-NO_3^- data for NO_3^- source identification, but in some cases effect size values are below 0 (e.g. manure and sewage), indicating negative contribution of $\delta^{18}O$-NO_3^- data for NO_3^- source identification. The overall grand mean effect size is 0.5 and the 95% confidence interval [0.13 to 0.83] does not overlap with zero. Thus, overall $\delta^{18}O$-NO_3^- data provide a significant contribution for NO_3^- source identification.

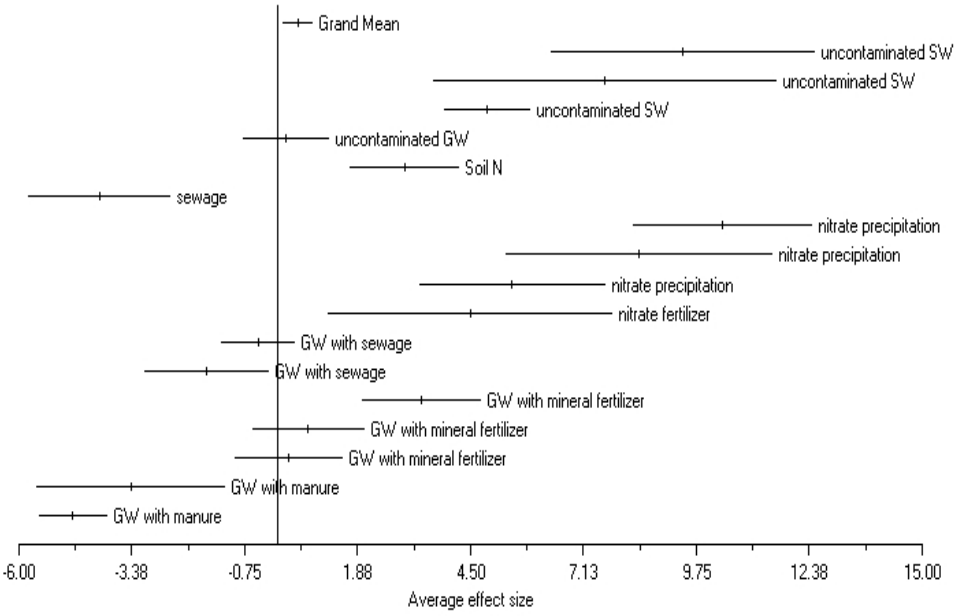


Figure 2-3: Average effect size and 95% confidence intervals of $\delta^{18}O$ on $\delta^{15}N$ for NO_3^- source identification. GW represents groundwater; SW represents surface water. GW with manure-Wassenaar (1995); Rock and Mayer (2002); GW with mineral fertilizer- Panno et al. (2001); Deutsch et al. (2006a); Moore et al. (2006); GW with sewage- Fukada et al. (2004); Moore et al. (2006); Nitrate fertilizer- Deutsch et al. (2005); Nitrate precipitation- Pardo et al. (2004); Deutsch et al. (2006a); Hales et al. (2007); Sewage- Aravena and Robertson (1998); Soil N- Hales et al. (2007); Uncontaminated GW- Moore et al. (2006); Uncontaminated SW- Karr et al. (2001); Pardo et al. (2004); Hales et al. (2007).

2.7 Factors influencing isotopic compositions of NO_3^- sources

Multiple NO_3^- sources from agricultural and urban activities contributing to surface- and groundwater and complex fractionations caused by multiple N-cycling processes change the original $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values, potentially biasing identification of NO_3^- contaminated sources (Kellman, 2005; Kendall, 1998).

Several soil N processes control the $\delta^{15}N$ values of both the NH_4^+ available to be nitrified and the formed NO_3^-. The ammonification process (organic-N $\rightarrow$ NH_4^+) results in a small fractionation (±1‰) between (soil) organic matter and (soil) NH_4^+ (Kendall, 1998). In contrast, the conversion of NH_4^+ to NO_2^- and NO_3^- is accompanied by marked N isotope fractionation effects, resulting in NO_3^- depleted in ^{15}N relative to the $\delta^{15}N$ of the initial NH_4^+, while the residual NH_4^+ is characterized by progressively increasing $\delta^{15}N$ values (Delwiche

and Steyn, 1970; Mariotti et al., 1981; Macko and Ostrom, 1994). The enrichment factors are between -12‰ to -29‰ (Kendall, 1998). During nitrification, the $\delta^{18}O$ of newly produced NO_3^- obviously does not depend on the isotope composition of the organic matter being mineralized, but depends on the $\delta^{18}O$ of H_2O and O_2 incorporation (see above). However, there is emerging evidence that O can exchange between H_2O and intermediate compounds of nitrification (Andersson et al., 1982; DiSpirito and Hooper, 1986; Kool et al., 2007). There are generally three steps in nitrification: firstly, NH_4^+ is oxidized to NH_2OH catalyzed by ammonia monooxygenase, in which O from O_2 is incorporated; secondly, NH_2OH is oxidized to NO_2^- catalyzed by hydroxylamine oxidoreductase, in which O from H_2O is incorporated; and finally, NO_2^- is oxidized to NO_3^- catalyzed by nitrite oxidoreductase with incorporation of O from H_2O, in which O can exchange between H_2O and NO_2^- or NO_3^- as a result of reversibility of this step (Kool et al., 2007).

Another possible process causing significant alterations of the isotopic composition of NO_3^- is microbial denitrification (heterotrophic denitrification), which occurs when O is limited and organic carbon is available (Knowles, 1982).:

$$4\,NO_3^- + 5CH_2O + 4H^+ \rightarrow 2N_2 + 5CO_2 + 7H_2O \qquad (2\text{-}6)$$

Chemo-autotrophic denitrification by bacteria such as Thiobacillus denitrificans, which oxidizes sulfur, can also be important (Batchelor and Lawrence 1978). Denitrification by pyrite oxidation is expressed as (Torrentó et al., 2010):

$$14\,NO_3^- + 5FeS_2 + 4H^+ \rightarrow 7\,N_2 + 10SO_4^{2-} + 5Fe^{2+} + 2H_2O \qquad (2\text{-}7)$$

Kinetic isotope effects during microbial denitrification are responsible for preferentially converting the lighter isotopes ^{14}N and ^{16}O to N_2 and N_2O, causing an enrichment of the heavy isotopes in the remaining NO_3^- (Mariotti et al., 1981; Mayer et al., 2002; Fukada et al., 2003). Nitrogen isotope enrichment factors of denitrification are within -40‰ to -5‰ (Hübner, 1986; Smith et al., 1991; Sebilo et al., 2003), and oxygen isotope enrichment factors are between -18‰ and -8‰ (Böttcher et al., 1990; Mengis et al., 1999; Fukada et al., 2003; Lehmann et al., 2003). Some studies reported that a linear relationship indicating an enrichment of ^{15}N relative to ^{18}O by a factor between 1.3:1 and 2.1:1 gives evidence for denitrification (Aravenaand and Robertson, 1998; Mengis et al., 1999; Fukada et al., 2003).

Ammonia volatilization is a process that causes a large enrichment of the heavier ^{15}N isotope in the residual NH_4^+. Volatilization involves equilibrium (reversible) reactions and kinetic (irreversible) reactions leading to fractionation. The equilibrium fractionation occurs between NH_4^+ and NH_3 in solution, and between aqueous and gaseous NH_3. The hydrolysis of urea (the main N form in animal waste) or NH_4^+ fertilizers causes a temporary increase in pH, which favors the NH_3 gas loss by volatilization. The kinetic fractionation is caused by the diffusive loss of ^{15}N-depleted NH_3, which results in NH_4^+ enriched in ^{15}N (Kendall, 1998).

Mixing of multiple NO_3^- sources is another bias that affects NO_3^- source identification. Although different NO_3^- sources have a distinctive isotopic composition, the mixture shows intermediate values. Moreover, NO_3^- sources in the environment rarely keep the initial composition as various fractionation processes may alter the initial composition before or after mixing (Kendall, 1998). This situation may give rise to incorrect calculations of the contribution of NO_3^- sources in mixing models and inaccurate results for NO_3^- source identification (Panno et al., 2006).

2.8 Source identification by combining isotope methods with hydrochemistry

Recently, researchers have attempted to link $\delta^{15}N$ and $\delta^{18}O$ values with land-use types or physico-chemical properties of water to identify NO_3^- sources (Karr et al., 2001; Mayer et al. 2002; Spruill et al. 2002; Mitchell et al. 2003; Kaushal et al., 2006). From three years of observation in rural areas, Choi et al. (2007) reported the following $\delta^{15}N$ values of groundwater contaminated with NO_3^- from different land-use activities: +4.5‰ to +8.5‰ for cropping areas with mineral fertilizer use, from +8.7‰ to +17.6‰ for farmland with animal manure application, and above +10‰ for residential areas with sewage discharge. Mayer et al. (2002) and Voss et al. (2006) found a significant positive linear relationship ($r^2 = 0.75$, $p = 0.001$, $n = 16$ in Mayer et al. (2002); $r^2 = 0.77$, $p < 0.001$, $n = 11$ in Voss et al. (2006)) between $\delta^{15}N$ values of riverine NO_3^- and percentages of agricultural and urban land-use. They showed that concentrations and N isotope values of riverine NO_3^- under agricultural plus urban land were markedly higher than in predominantly forested areas with soil N and atmospheric deposition as the main NO_3^- sources. Thus, evidence from the combination of isotopic composition values with land-use types indicates that anthropogenic activities are the predominant NO_3^- sources for contaminated surface- and groundwater.

The combination of concentration and isotopic composition of NO_3^- can also provide useful evidence for identification of NO_3^- sources (Kendall, 1998; Mayer et al. 2002; Pardo et al.

2004). Min et al. (2002) and Choi et al. (2007) found a correlation between increasing NO_3^- concentrations in groundwater and more positive $\delta^{15}N\text{-}NO_3^-$ values when manure was the predominant NO_3^- source. The contribution of manure raised the $\delta^{15}N\text{-}NO_3^-$ signatures above the background levels of soil-derived NO_3^- in groundwater. In contrast, Iqbal et al. (1997) observed a correlation between increasing NO_3^- concentrations in groundwater and more negative $\delta^{15}N\text{-}NO_3^-$ values when fertilizer was the predominant source. The contribution of fertilizers decreased the $\delta^{15}N\text{-}NO_3^-$ signatures below the background levels of soil-derived NO_3^- in groundwater. However, these correlations could only be functional in the case of a single NO_3^- contamination source and when no fractionation process occurs.

Linking $\delta^{15}N\text{-}NO_3^-$ values with concentrations of anions and cations in water also provides valuable information for identifying NO_3^- sources, as chemical constituents from anthropogenic sources are inclined to dominate those of natural sources (Spruill et al., 2002). Min et al. (2002) performed a study on NO_3^- sources from agricultural practices for riverside alluvial aquifers along the Nakdong River in Korea, and found that based on $\delta^{15}N\text{-}NO_3^-$ values and physico-chemical data (e.g. Ca^{2+}, Na^+, K^+, Cl^-, HCO_3^-, etc.) fertilizers, animal manure and human wastes were dominant NO_3^- sources. They reported that when the groundwater type changes from a Ca-HCO_3 type (background or less-polluted) to a Na(-K)-Cl type, local contamination by domestic sewage or manure might be responsible. Changes of the major anions from HCO_3^- to Cl^- are possibly caused by agricultural activities. Karr et al. (2001) investigated the effects of swine-manure contamination on groundwater using $\delta^{15}N\text{-}NO_3^-$ together with K^+, Cl^- and NO_3^- data. They pointed out that the source of Cl^- and K^+ enrichment was swine-manure and found positive correlations between NO_3^-, Cl^-, and K^+ in ground water. They determined the median $\delta^{15}N\text{-}NO_3^-$ of ground water to be +15.4‰, falling well in the range of manure (+5‰ and +25‰). Thus, demonstrating that swine-manure can be successfully traced. Spruill (2002) built a classification tree model using N isotopes, anions and cations to correctly classify five NO_3^- sources (fertilizer on crops, fertilizer on golf courses, irrigation spray from hog wastes, and leachate from poultry litter and septic systems) in groundwater and obtained a correct classification rate of more than 80%.

2.9 Boron isotope application for NO_3^- source identification

Apart from land-use types and hydrochemistry, boron isotopes (B) have also been used for the identification of NO_3^- sources in water (Bassett et al., 1995; Vengosh et al., 1999; Widory et al., 2004; Widory et al., 2005). Boron has two stable isotopes (^{11}B and ^{10}B) with natural

abundances of approximately 80% and 20%. Boron is highly soluble in aqueous solutions and is present as a minor or trace constituent in nearly all water types (Bassett et al., 1995; Chetelat and Gaillardet, 2005). Furthermore, B is not affected by transformation processes (Vengosh et al., 1994; Bassett et al., 1995). Boron has already been used in several studies as a tracer for sewage contamination in groundwater (Vengosh et al., 1994; Bassett et al., 1995; Eisenhut et al., 1997; Verstraeten et al., 2005) and surface water (Vengosh et al., 1999; Chetelat and Gaillardet, 2005). Values of $\delta^{11}B$ of sewage reported in the literature range from -7.7‰ to +12.9‰ (Gellenbeck, 1994; Vengosh et al., 1994; Bassett et al., 1995; Leenhouts et al., 1998; Vengosh et al., 1999; Widory et al., 2004; Widory et al., 2005; Seiler, 2005). Animal manure and mineral N fertilizers also contain B as a minor or trace element. The $\delta^{11}B$ values reported for different types of manure range from +6.9‰ to +42.1‰, and are predominately higher than the $\delta^{11}B$ values reported for fertilizers (ranging from +8‰ to +17‰) and the $\delta^{11}B$ values of sewage. These distinct $\delta^{11}B$ signatures have been used to discriminate between animal manure sources and sewage or, to a lesser extent, fertilizer sources (Komor, 1997; Widory et al., 2004; Widory et al., 2005). Widory et al. (2004, 2005) demonstrated the benefits of the combined use of $\delta^{15}N$ and $\delta^{11}B$ values to identify multiple NO_3^- sources (fertilizers, greenhouse discharges, sewage, hog, cattle and poultry manure) in groundwater in areas with different hydrogeological conditions in France. Seiler (2005) used $\delta^{15}N$ and $\delta^{18}O$ values combined with $\delta^{11}B$ values to discriminate between domestic wastewater and fertilizer contamination in groundwater in Nevada (USA).

The main processes which may affect the B isotopic composition are adsorption-desorption interactions with clay minerals, iron and aluminum oxide surfaces and organic matter (Palmer et al., 1987; Bassett, 1990; Goldberg et al., 2000; Yingkai and Lan, 2001; Lemarchand et al., 2005) and mineral precipitation (only in extremely saline environments) (Bassett et al., 1995).

2.10 Quantification of NO_3^- source inputs

Deutsch et al. (2006a) successfully used the dual isotope approach to quantify three NO_3^- sources (water from artificially drained agricultural soils, groundwater and atmospheric deposition). They also determined the contribution of each NO_3^- source to a river sub- basin in Germany based on a mass-balance mixing model (Phillips and Koch, 2002). Results from the mixing model indicated that NO_3^- from the drainage water contributed 86% and the other two NO_3^- sources contributed 11% (groundwater) and 3% (atmospheric deposition) respectively. The dual isotope approach was also applied to quantify NO_3^- inputs into 12 Baltic rivers by

Voss et al. (2006). Similarly, a mass-balance mixing model was used to quantify three major N source contributions, which are sewage, atmospheric deposition and pristine soils.

Normally, a basic mass-balance mixing model based on two isotopes and three sources is expressed as follows:

$$\delta^{15}N_M = f_1\delta^{15}N_1 + f_2\delta^{15}N_2 + f_3\delta^{15}N_3 \quad (2\text{-}8)$$

$$\delta^{18}O_M = f_1\delta^{18}O_1 + f_2\delta^{18}O_2 + f_3\delta^{18}O_3 \quad (2\text{-}9)$$

$$1 = f_1 + f_2 + f_3 \quad (2\text{-}10)$$

where $\delta^{15}N_M$ and $\delta^{18}O_M$ are the NO_3^- isotope values from the mixture, and the subscripts 1, 2 and 3 represent any of the three NO_3^- sources; f is defined as the proportional contribution of the respective source.

However, Moore and Semmens (2008) mentioned that this mixing model does not take into account the incorporation of several substantial sources of uncertainty. The first source of uncertainty is temporal and spatial variability in $\delta^{15}N$- and $\delta^{18}O$-NO_3^- in NO_3^-; the second source of uncertainty is isotope fractionation during denitrification; and the third is that too many NO_3^- sources contribute to the mixture so that the above equations cannot be solved exactly. To fully incorporate those sources of uncertainty, Parnell and Jackson (2008) have developed and implemented a stable isotope mixing model called SIAR, which is a recently released open source software package that runs in R statistical computing program. This model has been successfully applied in food-web analysis, such as estimates of prey (source) contributions to the predator (mixture). SIAR uses a Bayesian framework to establish a logical prior distribution based on Dirichlet distribution (Evans et al., 2000) for estimating possible proportional source contribution, and then to determine the probability distribution for the proportional contribution of each source to the mixture. By defining a set of N mixture measurements on J isotopes with K source contributors, the mixing model which takes the above mentioned uncertainties into account can be expressed as follows (Parnell et al., 2010):

$$X_{ij} = \sum_{k=1}^{K} p_k(S_{jk} + c_{jk}) + \varepsilon_{ij} \quad (2\text{-}11)$$

$$S_{jk} \sim N\left(\mu_{jk}, \omega_{jk}^2\right)$$

$$c_{jk} \sim N\left(\lambda_{jk}, \tau_{jk}^2\right)$$

$$\varepsilon_{ij} \sim N\left(0, \sigma_j^2\right)$$

where X_{ij} is the isotope value j of the mixture i, in which $i = 1, 2, 3,\ldots, N$ and $j = 1, 2, 3,\ldots, J$; S_{jk} is the source value k on isotope j ($k = 1, 2, 3,\ldots, K$) and is normally distributed with mean μ_{jk} and standard deviation ω_{jk}; p_k is the proportion of source k, which needs to be estimated by the SIAR model; c_{jk} is the fractionation factor for isotope j on source k and is normally distributed with mean λ_{jk} and standard deviation τ_{jk}; and ε_{ij} is the residual error representing the additional unquantified variation between individual mixtures and is normally distributed with mean 0 and standard deviation σ_j. The detailed description of this model can be found in Moore and Semmens (2008) and Jackson et al. (2009) and Parnell et al. (2010).

2.11 Analytical techniques for the determination of $\delta^{15}N$- and $\delta^{18}O$-NO_3^-

Precise, accurate, but also inexpensive and fast analysis of NO_3^- for both $\delta^{15}N$ and $\delta^{18}O$ is needed for improved NO_3^- source identification, quantification and uncertainty assessment.
In recent years, the so called "ion-exchange" or "$AgNO_3$ method" for both $\delta^{15}N$-NO_3^- and $\delta^{18}O$-NO_3^- analysis has been developed by Chang et al. (1999) and Silva et al. (2000). This method is used to concentrate and purify NO_3^- in water samples for simultaneous ^{15}N and ^{18}O determination. Briefly, NO_3^- is purified and concentrated by passing samples through cation and then anion exchange resin columns. NO_3^- is eluted using hydrochloric acid, neutralized with silver oxide and then filtered to remove the AgCl. For accurate $\delta^{18}O$-NO_3^- analysis, all non-nitrate oxygen-bearing anions (e.g. SO_4^{2-}, CO_3^{2-}, and PO_4^{3-}) are removed from the sample by adding $BaCl_2$ to the $AgNO_3$ solution and then the precipitate is filtered out. The filtered solution is then passed through a cation exchange resin to remove excess Ba^{2+} ions and re-neutralized with Ag_2O. The resulting solution can be freeze-dried (Silva et al., 2000) or oven-dried (Fukada et al., 2003) to produce $AgNO_3$ salts. $\delta^{15}N$ analysis of the prepared $AgNO_3$ can be conducted by conversion to N_2 gas for IRMS analysis via mixing $AgNO_3$ with a CuO, Cu wire and CaO catalyst and combusting in a sealed tube at 850˚C (Kendall and Grim, 1990). $\delta^{18}O$ analysis can be conducted using the combustion method which generates CO_2 by adding finely ground spectrographic graphite or by using pyrolysis systems that can generate CO by addition of graphite. Alternatively, $\delta^{15}N$ and $\delta^{18}O$ can be simultaneously analyzed via TC/EA-IRMS (thermal conversion/elemental analyzer-isotope ratio mass spectrometer). Prepared $AgNO_3$ samples are converted (N_2 and CO) at 1400˚C in a molybdenum-lined,

aluminum oxide combustion tube which is filled with glassy carbon and topped with a glassy carbon crucible. Upon pyrolysis of the $AgNO_3$, the reaction gases are separated via a 1m E3030 GC column (Elemental Microanalysis) and analyzed via IRMS. The "ion-exchange method" has the following advantages: (1) the concentration of NO_3^- from water samples onto anion exchange resin columns can be accomplished in the field and is convenient for transporting and storing water samples for NO_3^- isotopic analysis; (2) there is minimal isotopic fractionation for NO_3^- stored on anion exchange columns (Silva et al., 2000); and (3) the technique achieves a high level of sensitivity. The disadvantages of the "ion-exchange method" can be summarized as follows: (1) the sample preparation procedure is relatively labor-intensive (3-5 days for sample preparation) and cost-intensive (up to 60 Euro per sample just for consumables only); (2) high concentrations of anions (e.g. Cl^-, SO_4^{2-}, DOC, etc) in water samples can interfere with the adsorption of NO_3^- onto anion exchange resins; and (3) target sample size of 100-200 µmol of NO_3^- for optimal analysis requires large sample volumes for low NO_3^- concentration samples.

Another sample preparation technique for both $\delta^{15}N\text{-}NO_3^-$ and $\delta^{18}O\text{-}NO_3^-$ determination is the so called "bacterial denitrification method" (Sigman et al., 2001; Casciotti et al., 2002). This method allows for the simultaneous determination of $\delta^{15}N$ and $\delta^{18}O$ of N_2O produced from the conversion of NO_3^- by denitrifying bacteria which naturally lack N_2O-reductase activity. In brief, bacterial cultures are grown for 6-10 days in amended tryptic soy broth (TSB), divided into centrifuge tubes of 40 mL aliquots and centrifuged. After centrifugation, the supernatant is decanted, reserved and 4 mL of the TSB is pipetted back into the tubes to obtain a 10-fold concentration of bacteria. These tubes are then vortexed to ensure homogenized cultures and then transferred as 2 x 2 mL aliquots into 20 mL headspace vials. The vials are crimp-sealed with Teflon-backed silicone septa. To ensure anaerobic conditions, a reduced blank effect and removal of N_2O produced prior to sample injection, the headspace vials are purged with N_2 gas for 3 hours. Samples of dissolved NO_3^- (100 nmol) are then injected into the headspace vials and are incubated overnight to allow for complete conversion of NO_3^- to N_2O. The next day, 0.1 mL of 10N NaOH is injected into the headspace vials to stop bacterial activity and to scrub any CO_2 gas in the vial which can interfere with the N_2O measurement. $\delta^{15}N$ and $\delta^{18}O$ analysis of the produced N_2O can be conducted by (1) extracting N_2O offline via freezing it out in a vacuum line (cryo-concentration) and analyzing by IRMS (Sigman et al., 2001), or (2) extracting N_2O online via an autosampler injection system and using Nafion as an additional water trap, and then analyzing by IRMS (Casciotti et al., 2002). The "bacterial denitrification method" has the following advantages: (1) the preparation

procedure is less labor-intensive (2-3 days for sample preparation) and less expensive (up to 5 Euro per sample for consumables); (2) the technique requires a smaller sample size (three orders of magnitude smaller than the "ion-exchange method") and allows for the analysis of low NO_3^- concentration samples; and (3) the method achieves a high level of sensitivity. The disadvantages of the "bacterial denitrification method" are summarized as follows: (1) bacterial growth takes a longer time of about 10-12 days (from Petri plate to media bottles); (2) bacterial cultivation is potentially affected by toxicity of the sample (e.g. antibiotics, heavy metal, pesticides, etc), which is difficult to predict; (3) the presence of NO_2^- in water samples can bias the isotopic composition of product N_2O; and (4) mathematical approaches are needed to quantify and correct for both fractionation of oxygen isotopes during oxygen atom loss in the reaction process and exchange of oxygen atoms in nitrite and nitric oxide with water that are inherent to the denitrifier method for $\delta^{18}O$ analysis.

2.12 Conclusions

Various potential NO_3^- sources have distinct $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values and the dual isotope approach is a powerful tool to identify NO_3^- sources in contaminated water. Furthermore, incorporation of hydrochemistry may help to confirm NO_3^- source identification. However, complex fractionation processes and admixture from multiple NO_3^- sources usually make it difficult to correctly identify the NO_3^- source contribution. These problems may be eliminated by combining $\delta^{15}N$- and $\delta^{18}O$-NO_3^- with co-migrating discriminators of NO_3^- sources. For example, boron (B) isotopes, which are present as a minor or trace constituent in nearly all water types, are not affected by biogeochemical transformation processes. Meanwhile, quantification and uncertainty assessment of a variety of NO_3^- source inputs via Bayesian framework estimation could provide powerful evidence for NO_3^- source contribution. In addition, the "ion-exchange method" and the "bacterial denitrification method" are frequently used for $\delta^{15}N$- and $\delta^{18}O$-NO_3^- measurement, but only a limited number of laboratories worldwide are capable to carry out these analyses. Hence, precise, accurate, efficient and broadly applicable isotope techniques need to be explored before NO_3^- source quantification using the isotope mixing model operating in a Bayesian framework can be implemented in large scale water monitoring programs.

Chapter 3:

Sampling point selection for isotope monitoring based on a decision tree model for nitrate source identification in surface water in Flanders

3.1 Abstract

The 47 sampling points were selected by experts from the MAP (Manure Action Plan) monitoring network for NO_3^- source identification in surface water in Flanders. The sampling points were classified into 5 different NO_3^- source classes defined by experts: agriculture (class A, 7 sampling points); agriculture with groundwater compensation (class AGC, 15 sampling points); a combination of agriculture with horticulture (class AH, 6 sampling points); greenhouses in an agricultural area (class G, 11 sampling points) and households (class H, 8 sampling points). The purpose of this chapter is to use physicochemical data in a decision tree model to evaluate the classification results of the 47 sampling points. Based on the outcome of this decision tree, representative sampling points for isotope monitoring have been selected for retrieving the expert classification in subsequent chapters.

The decision tree model obtained an overall 82% CCI (percentage of correctly classified instance). Classes A, AGC, G, and H were classified well (more than 80% CCI compared to expert classification), while class AH had the lowest percentage (58% CCI). Sampling points in class AH demonstrated a low reconstrutability given by physico-chemical data compared to the expert knowledge. The seasonal difference of the CCI for individual sampling points between summer and winter varied widely, from 1 to 21%, which indicated that season has an effect on the decision tree model performance. Finally, thirty sampling points (6 sampling points per class) were selected based on using the decision tree model output.

3.2 Introduction

The implementation of the Nitrate Directive (EC, 2002) in Europe established a detailed framework for the protection of waters due to NO_3^- pollution from agricultural sources. While this directive imposed maximum allowable NO_3^- concentrations (50 mg NO_3^- L^{-1}) and initialized an extensive NO_3^- monitoring campaign, some regions continue to show elevated if not increasing NO_3^- levels (VMM, 2005).

In Chapter 1, the status of NO_3^- contamination in surface water in Flanders was shown in Figure 1-4. NO_3^- contamination was reduced between 1999 and 2003, resulting from the effectiveness of the MAP action. After 2003, the reduction in NO_3^- concentration seems to stagnate on average in Flanders. Thus, it is essential to find out the reasons for the stagnation and better knowledge of NO_3^- sources in the different basins is required. Therefore, we proposed to develop a technique identifying NO_3^- sources via isotopic fingerprinting.

Experts from VMM roughly classified the monitoring sampling points of MAP into five NO_3^- source classes: classes A (agriculture), AGC (agriculture with ground water compensation), AH (a combination of agriculture with horticulture), G (greenhouses in an agricultural area), and H (households). Each sampling point has been monitored for 10 physico-chemical parameters (temperature (T), electrical conductivity (EC 20), pH, dissolved oxygen (DO), oxygen saturation (O_{2sat}) , chloride (Cl^-), ammonium (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-) and phosphate (PO_4^{3-})) since at least 2002. Thus, the time series of physico-chemical data of these 47 sampling points comprised a relatively large data set, requiring a model to extract key information, which allowed us to evaluate the 47 sampling points.

A decision tree model has been developed in this study using the data of 47 sampling points. The basic idea behind a decision tree model is to create a hierarchical tree based on a data set of known classes, and then the resulting tree is used to predict classes from another independent data set having the same parameters but unknown classes. The advantage of a decision tree model is that it is graphical and produces a visual output and is easily interpreted. This model guides users through a series of if-then statements from the beginning of the tree through a series of subgroups to the final group classification.

The purpose of this chapter is to explore a decision tree model to evaluate the reconstrutability given by physico-chemical data compared to the expert knowledge and select representative sampling points for isotope monitoring to retrieve of expert classification in later chapters.

3.3 Material and methods

3.3.1 Description of the sites

The selected 47 sampling points were distributed over different basins in Flanders (Figure 3-1).

Five NO_3^- source classes were defined as: agriculture (class A), agriculture with groundwater compensation (class AGC), a combination of agriculture with horticulture (class AH), greenhouses in an agricultural area (class G) and households (class H). The coordinates of the sampling locations and mean NO_3^- concentrations of these sampling points during a monitoring period from 2002 - 2009 are shown in Table 3-1.

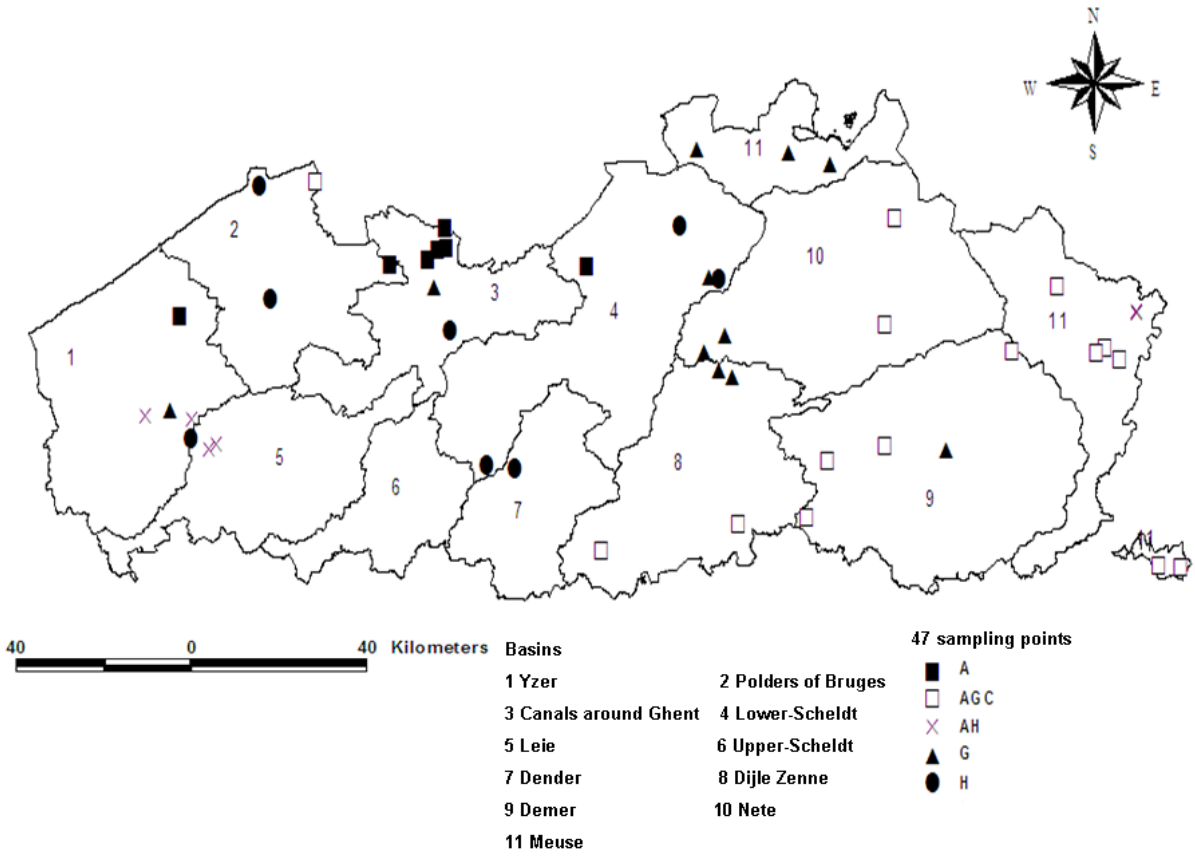


Figure 3-1: Location and classification of 47 sampling points in Flanders. A represents agriculture; AGH represents agriculture with groundwater compensation; AH represents a combination of agriculture with horticulture; G represents greenhouses in an agricultural area; H represents households.

Table 3-1: Summary information regarding 47 MAP sampling points.

VMM Number	NO_3^- class	$NO_3^{-\$}$ Concentration (mg N L^{-1})	Basin	Coordinates
13000	A	2.9 ± 3.1	Canals around Ghent	51°16'12" N, 3°45'10" E
13500	A	1.8 ± 2.1	Canals around Ghent	51°14'21" N, 3°45'20" E
17000	A	3.9 ± 3.3	Canals around Ghent	51°14'8" N, 3°43'45" E
19000	A	5.7 ± 3.7	Canals around Ghent	51°13'6" N, 3°42'17" E
24050	A	5.6 ± 4.3	Polders of Bruges	51°12'44" N, 3°35'36" E
194660	A	3.0 ± 3.8	Lower-Scheldt	51°12'47" N, 4°10'2" E
861110	A	14.2 ± 11.0	Yzer	51°7'36" N, 2°58'45" E
26720	AGC	1.1 ± 1.2	Polders of Bruges	51°20'28" N, 3°22'16" E
107750	AGC	22.4 ± 2.7	Meuse	51°10'28" N, 5°32'12" E
119300	AGC	3.1 ± 1.1	Meuse	51°4'41" N, 5°40'18" E
136010	AGC	4.3 ± 0.5	Meuse	51°4'7" N, 5°38'52" E
137550	AGC	0.8 ± 0.6	Meuse	51°3'29" N, 5°42'54" E
149100	AGC	11.3 ± 2.1	Meuse	50°44'3" N, 5°49'18" E
153400	AGC	12.6 ± 4.3	Meuse	50°43'51" N, 5°52'51" E
307100	AGC	1.4 ± 0.5	Nete	51°17'10" N, 5°3'49" E
328520	AGC	0.7 ± 0.5	Nete	51°7'15" N, 5°2'11" E
365520	AGC	6.5 ± 2.0	Dijle Zenne	50°45'57" N, 4°12'36" E
408760	AGC	9.8 ± 2.8	Demer	50°54'17" N, 4°51'48" E
426520	AGC	11.4 ± 1.0	Demer	50°49'N, 4°48'4" E
426605	AGC	11.9 ± 3.1	Demer	50°55'48" N, 5°1'59" E
453920	AGC	20.4 ± 5.0	Demer	51°4'29" N, 5°24'13" E
485550	AGC	8.2 ± 1.2	Dijle Zenne	50°48'25" N, 4°36'21" E
130300	AH	19.3 ± 12.8	Meuse	51°7'59" N, 5°46'12" E
130350	AH	29.8 ± 14.4	Meuse	51°8'4" N, 5°46'11" E
624515	AH	18.1 ± 8.4	Leie	50°55'31" N, 3°5'36" E
624550	AH	13.5 ± 8.7	Leie	50°55'6" N, 3°4'16" E
926100	AH	19.1 ± 11.3	Yzer	50°57'54" N, 3°1'12" E
938210	AH	21.3 ± 12.3	Yzer	50°57'59" N, 2°53'8" E
35450	G	9.5 ± 6.6	Canals around Ghent	51°10'34" N, 3°43'21" E
65100	G	6.9 ± 8.5	Meuse	51°23'40" N, 4°29'11" E
82870	G	12.9 ± 18.1	Meuse	51°23'18" N, 4°45'17" E
83900	G	8.9 ± 12.8	Meuse	51°22'17" N, 4°52'41" E
190220	G	9.6 ± 11.7	Lower-Scheldt	51°11'41" N, 4°31'27" E
263100	G	41.4 ± 37.9	Nete	51°4'43" N, 4°30'35" E
269030	G	18.2 ± 17.9	Nete	51°6'9" N, 4°34'9" E
376220	G	28.5 ± 29.1	Dijle Zenne	51°2'59" N, 4°33'12" E
376240	G	44.4 ± 36.8	Dijle Zenne	51°2'17" N, 4°35'26" E
449890	G	44.2 ± 40.0	Demer	50°55'12" N, 5°12'36" E
941000	G	17.9 ± 10.1	Yzer	50°58'35" N, 2°57'34" E
6009	H	0.4 ± 0.2	Polders of Bruges	51°19'51" N, 3°12'41" E
57500	H	1.1 ± 2.2	Canals around Ghent	51°6'28" N, 3°46'6" E
183700	H	0.4 ± 0.4	Lower-Scheldt	51°16'31" N, 4°26'17" E
190260	H	0.7 ± 1.3	Lower-Scheldt	51°11'25" N, 4°33'11" E
520400	H	3.0 ± 2.0	Dender	50°53'38" N, 3°57'25" E
546000	H	3.3 ± 1.5	Lower-Scheldt	50°53'51" N, 3°52'29" E
629100	H	3.6 ± 4.7	Leie	50°55'58" N, 3°1'14" E
883560	H	1.2 ± 1.3	Polders of Bruges	51°9'18" N, 3°14'43" E

$ average nitrate concentration between 2002 and 2009 plus minus standard deviation

The historical mean NO_3^- concentrations were relatively high in classes AH and G, intermediate in classes A and AGC and low in class H. It has been suggested that in agricultural areas, the NO_3^- concentrations were mainly determined by mineral fertilizers and manure application. In household area, the NO_3^- concentrations were largely influenced by the habitant intensity and locations of sewage discharging points.

3.3.2 Data set development

A data set has been established according to the sampling point's code in the MAP network, sampling time and measurement of 10 physico-chemical parameters (temperature (T), electrical conductivity (EC 20), pH, dissolved oxygen (DO), oxygen saturation (O_{2sat}) , chloride (Cl^-), ammonium (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-) and phosphate (PO_4^{3-})). A fragment of the data set is shown in Table 3-2. Before discussing the question of how classification methods operate, it is important to give some useful mathematical definitions used in this thesis: (1) "data" is a single measurement of one parameter; (2) an "instance" is a combination of measurements of all parameters at one sampling time, i.e. one row in Table 3-2; (3) an "attribute" is all the measurements of one parameter for all time periods, i.e. one column in Table 3-2; and (4) "target class" is one of the five NO_3^- source classes.

The distribution of available instances over the five classes is shown in Figure 3-2. The number of instances in each class was equal to the sum of all events, all measurements of all parameters for all sampling points in a class. The total number of instances for the 47 sampling points was 3928. Class A and class G had most instances, which were related to the number of sampling points in the class and the sampling and measurement frequency. The number of instances in class AGC was a little less than the first two classes, but higher than class AH and class H.

Table 3-2: A fragment of VMM data set for decision tree model development.

Code number - year - month	Cl^- (mg L^{-1})	EC 20 (µS cm^{-1})	NH_4^+ (mg N L^{-1})	NO_2^- (mg N L^{-1})	NO_3^- (mg N L^{-1})	O_2 (mg L^{-1})	O_{2sat} (%)	PO_4^{3-} (mg P L^{-1})	pH	T (°C)	Target class
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
17000 1990 12	86	1005	1.5	0.2	12.5	7			7.6	6.1	A
17000 1991 01	95	1030	1.1	0.2	16.8	11.2		0.04	7.6	0.1	A
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
426520 2004 04	42	676	0.3	0	10	11	96	0.02	7.6	8.8	AGC
426520 2004 05		691			11.6	8.5	81		7.7	11.6	AGC
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
130300 2000 03	20	479	0.2	0	23	9.4	78	0.02	6.5	6.9	AH
130300 2000 04		375		0.2	6	6.9	72		6.6	17	AH
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
35450 2006 01	46	638	0.7	0.1	12	8.3	63	0.01	7.5	3.9	G
35450 2006 02	42	605	0.6	0.1	11	8.9	67	0.02	7.5	4.4	G
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
57500 2005 02	86	694	15	0	0.1	4.2	36	1.2	7.6	8	H
57500 2005 03		1185				2.8	23		7.5	8.1	H
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.

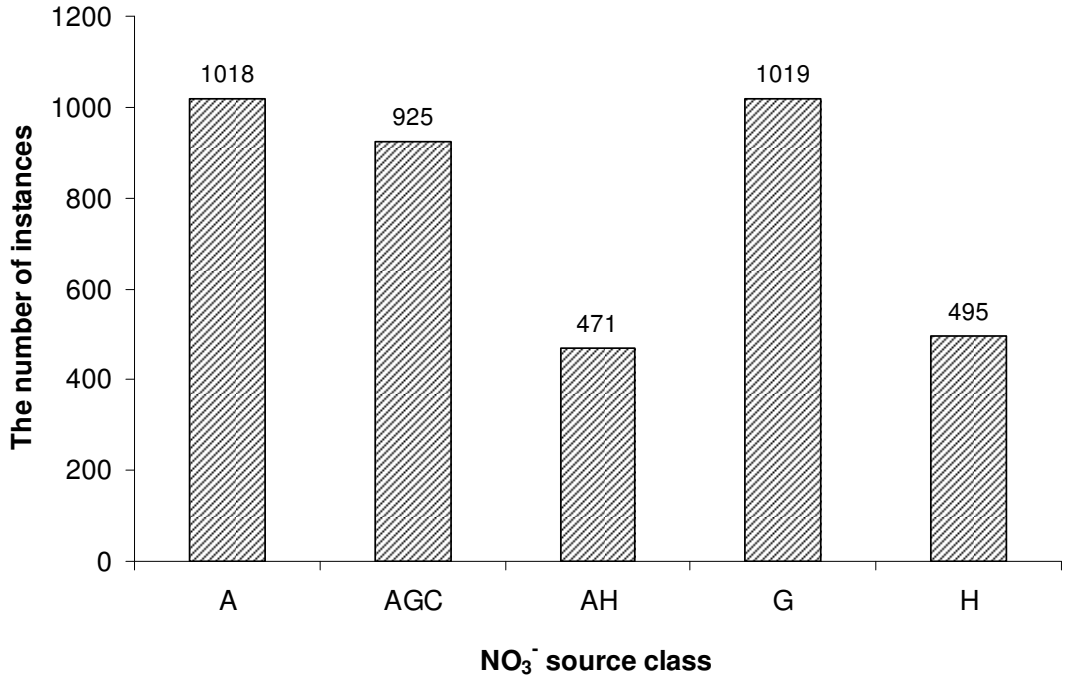


Figure 3-2: The distribution of instances over five source classes. A represents agriculture; AGH represents agriculture with groundwater compensation; AH represents a combination of agriculture with horticulture; G represents greenhouses in an agricultural area; H represents households.

A major concern in this study was the issue of missing data, as the physico-chemical parameters were not always measured at each sampling event. For example, for one instance in Table 3-2, with code number 426520 2004 05, only part of the physico-chemical parameters were measured in May 2004. Those unmeasured parameters were represented by blanks in the data set. Missing data occur for all the physico-chemical parameters which were measured, and the missing data percentage of one parameter is equal to the number of missing data divided by the theoretical total number of instances (3928). Results are listed in Table 3-3.

Table 3-3: Percentage of missing data each parameter.

	pH	T	EC 20	O_2	NO_3^-	O_{2sat}	NO_2^-	NH_4^+	PO_4^{3-}	Cl^-
Missing % for class A	4.8	4.8	6.9	6.0	2.3	28.5	21.3	31.5	31.9	42.1
Missing % for class AGC	5.4	5.4	5.6	7.2	5.9	14.6	39.5	44.5	47.8	52.7
Missing % for class AH	8.1	7.0	8.0	8.4	1.3	25.2	23.7	39.7	34.1	44.7
Missing % for class G	5.1	6.1	5.1	5.8	4.9	17.2	41.7	47.7	48.7	50.4
Missing % for class H	0.7	2.1	1.8	1.0	14.8	24.8	20.8	33.6	36.2	33.4
Mean missing %	4.8	5.1	5.6	5.7	5.8	22.1	29.4	39.4	39.7	44.7

Missing % represents missing percentage

The mean percentage of missing data varied widely from 4.8% to 44.7%. Some physico-chemical parameters (pH, T, EC20, O_2 and NO_3^-) have relatively low missing data (< 6%), while other parameters (O_{2sat}, NO_2, NH_4^+, PO_4^{3-} and Cl^-) have missing data above 20%. Furthermore, missing data percentages varied among different classes, but showed the same trend of mean missing percentages for all classes. The solution to deal with missing data in this study was imputation (a process where a reasonable alternative value is substituted for one that is missing) (Fielding et al., 2006). To get a complete data set, missing data was replaced with the average value for all valid data of that specific parameter for the corresponding sampling point. This method is easy to use, widely applied and allows to use a complete data set in the decision tree model.

3.3.3 Decision tree model

The basic algorithm for decision tree induction is a greedy algorithm that constructs decision trees in a top-down recursive divide-and-conquer manner. A decision tree model begins by separating the initial group composed of all instances (termed a parent node) into two sub-groups (termed child nodes) by examining all possible attributes and then selecting the best splitting attribute. The two resulting child groups are now the new parent nodes. They are then split into two more child nodes. This procedure continues until all instances are classified. The WEKA 3.4.10 software has been used to explore a decision tree. This software uses J4.8 algorithm, which is WEKA's implementation of the C4.5 decision tree learner (Quinlan, 1993). C4.5 is a decision tree program that is the machine learning workhorse most widely used in practice to date.

The most important criterion in decision tree construction is the split selection. If we had a measure of the purity of each node, we could choose the attribute that produces the purest child nodes. As a consequence, a critical step is the computation of the purity. Several purity functions can be chosen, yet the one mostly used is information gain. This function relies on the concept of entropy, which has to be understood here in the sense of information theory (Shannon, 1948). A given set T can be regarded as a distribution over s class labels, and its entropy can be calculated as

$$info(T) = -\sum_{i=1}^{s} p(c_i)\log_2 p(c_i) \tag{3-1}$$

where $p(c_i)$ denotes the proportion of T belonging to class c_i. The information gain ($gain(X)$) represents the expected reduction in entropy when splitting on attribute X and is calculated as:

$$gain(X) = info(T) - info(T|X) = info(T) - \sum_{j \in V(X)} \frac{|T_j|}{|T|} info(T_j) \tag{3-2}$$

where $V(X)$ denotes the number of possible values for attribute X and T_j is the subset of T for which attribute X has value j. The best attribute to be used as a decision criterion is the one for which the information gain is maximized (as maximizing the information gain minimizes the impurity).

An important aspect when finishing a decision tree model construction is to evaluate the performance of the model. To get a reliable result, a commonly used method is 10-fold cross validation (Figure 3-3).

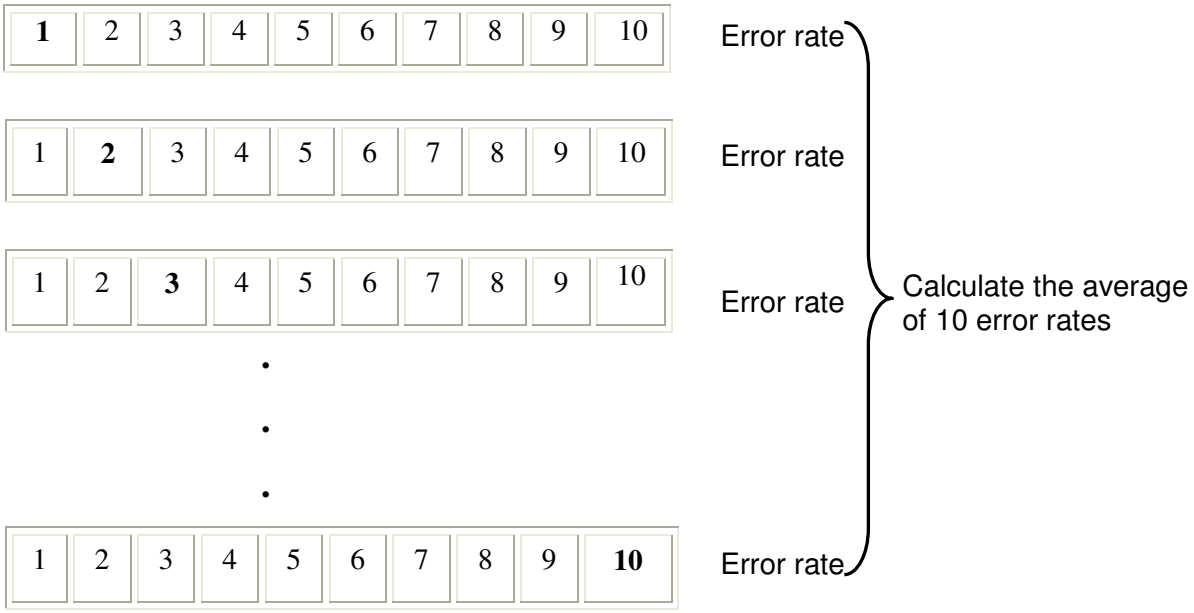


Figure 3-3: Scheme for 10-fold cross-validation procedure. The full data set is randomly divided into 10 parts and the testing parts are the numbers in bold.

Ten-fold cross-validation is a method by which the data set is divided randomly into 10 equal parts, in which each class is represented in approximately the same proportion as in the full data set. Nine parts are used for model learning and one part is used for model testing to calculate the error rate. Finally, the 10 error rates are averaged (Figure 3-3). The error rate is equal to one minus the ratio of the number of correct classification instances divided by the total number of classification instances.

3.4 Results and discussion

3.4.1 Decision tree model performance on five NO_3^- source classes

The decision tree model which was built based on a complete data set composed of 47 sampling points with 10 physico-chemical parameters had 247 leaves and the tree size is 493 (tree size is too big too show). There were 3230 instances out of 3928 classified according to the classes given by the experts. The rest (698 instances) were classified differently. Thus, the CCI compared to expert classification was 82%. Figure 3-4 displayed the correctly and incorrectly classified percentages for the five NO_3^- source classes compared to expert classification.

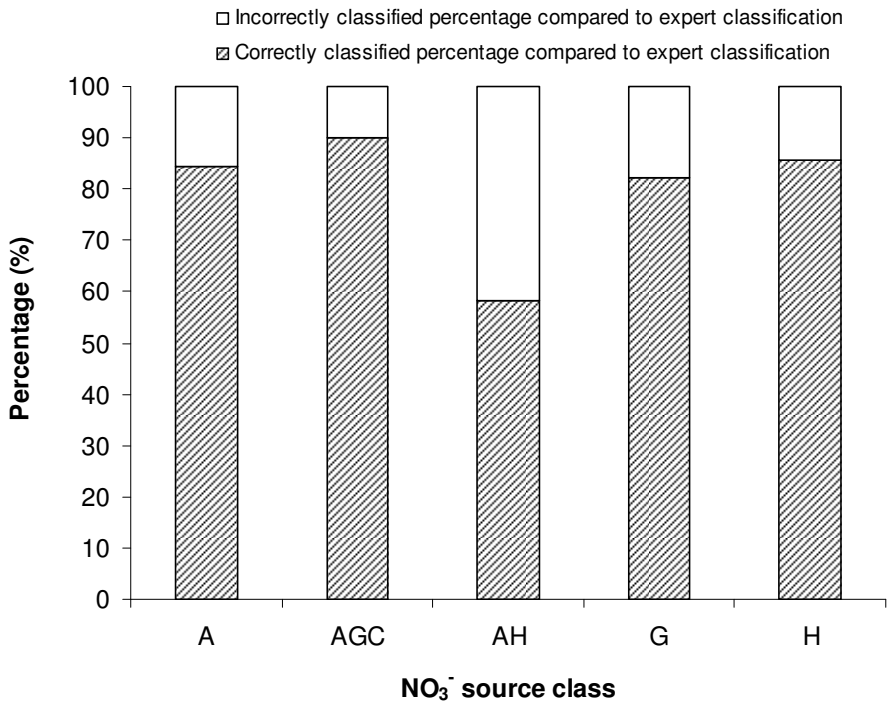


Figure 3-4: The percentage of correctly and incorrectly classified for five NO_3^- source classes compared to expert classification. A represents agriculture; AGH represents agriculture with groundwater compensation; AH represents a combination of agriculture with horticulture; G represents greenhouses in an agricultural area; H represents households.

It is clear that classes A, AGC, G, and H were classified well (more than 80% CCI compared to expert classes), which suggested a high reconstrutability given by physico-chemical data compared to the expert knowledge. Class AH showed the lowest CCI (58%), which implied the low reconstrutability given by physico-chemical data compared to the expert knowledge. Class AH had nearly the same number of instances as class H (Figure 3-2) and similar missing data percentages (Table 3-3), but it obtained a much lower classification result. It seems that the number of instances does not affect the classification result. The true reasons leading to high misclassification in class AH are not clear and need to be further studied by isotope monitoring in the following chapters. It is of interest to know the distribution of the five NO_3^- source classes after classification (Figure 3-5).

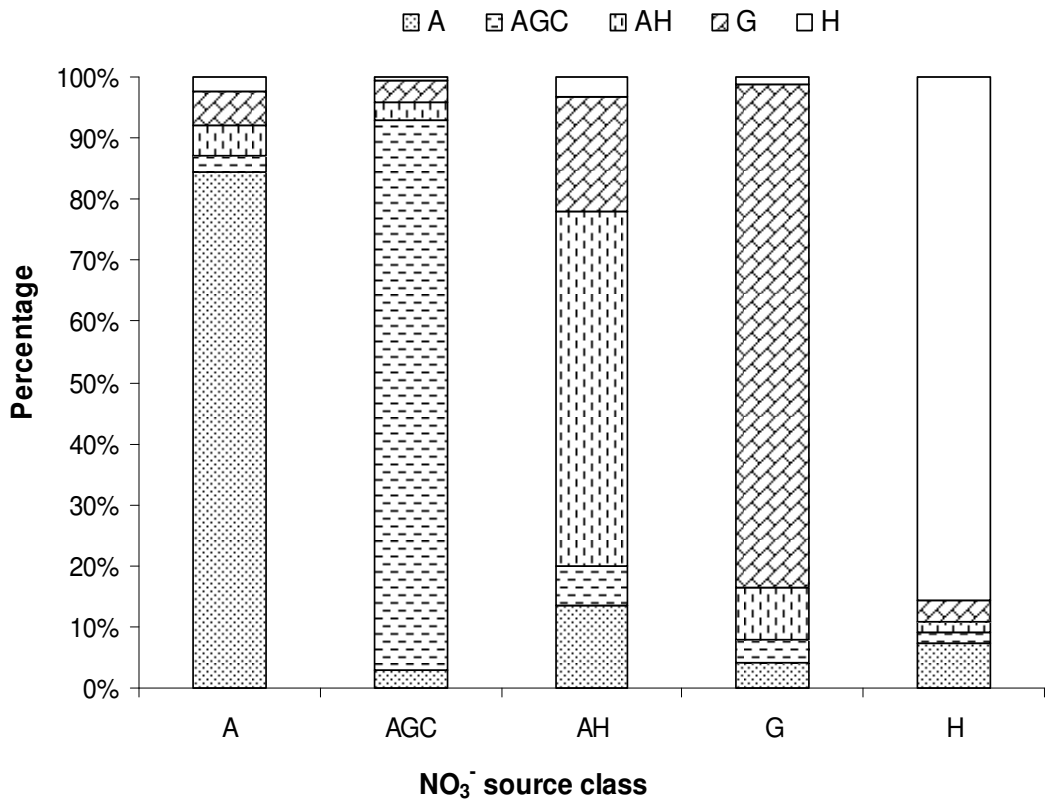


Figure 3-5: Distribution of all NO_3^- classes after classification in one particular class. A represents agriculture; AGH represents agriculture with groundwater compensation; AH represents a combination of agriculture with horticulture; G represents greenhouses in an agricultural area; H represents households.

It is obvious that the classes A, AGC, G and H with a high CCI had low CCI from other classes. However, in class AH with a low CCI, classes A and G occupied a relatively large percentage. Since class AH was the combination of agriculture and horticulture, the class label was not as pure as other classes. Data of some sampling points defined in this class may behave as the patterns in other classes.

3.4.2 Decision tree model performance on 47 sampling points

Table 3-4 display the CCI per sampling point and the evaluation of classification results for winter and summer (winter is from October to next March and summer is from April to September). The differently classified instances are the instances which were classified differently from expert classification, applying 10-fold cross validation.

As shown in Table 3-4, most sampling points had differently classified instances both in winter and summer, except for 100% similarity compared to expert classification of three sampling points in summer (sampling points 137550, 183700, and 328520) and one in winter (sampling point 107750). Although the overall mean CCI of the 47 sampling points in summer and winter were equal (82%), the difference of the CCI for individual sampling points between summer and winter varied widely, from 1 to 21%, which indicated that season has an effect on decision tree model performance. The mean CCI of the 47 sampling points ranged from 51 to 96%, and nearly three quarters of the sampling points obtained more than 80% similarity compared to expert classification. However, sampling points in class AH had relatively low classification results (less or equal to 65%). Sampling points in this class implied the low reconstrutability given by physico-chemical data compared to the expert knowledge. Thus, it is crucial to retrieve expert classification by an independent method, which would provide more reasonable evidence on the classification of the sampling points, not mainly based on different land uses defined by experts. Therefore, a dual isotope approach ($\delta^{15}N$- and $\delta^{18}O$-NO_3^-) will be applied in this thesis for discriminating NO_3^- pollution sources. Furthermore, it is also interesting to find the mechanism behind the seasonal differences.

Table 3-4: Evaluation of CCI for 47 sampling points compared to expert classification.

VMM number	Expert classification	Total instance	Total instance in summer	Incorrectly classified instance in summer	Total instance in winter	Incorrectly classified instance in winter	CCI in summer (%)	CCI in winter (%)	Mean CCI (%)
13000	A	214	103	13	111	15	87	86	87
13500	A	169	81	11	88	15	86	83	85
17000	A	203	96	24	107	20	75	81	78
19000	A	197	92	13	105	14	86	87	86
24050	A	120	55	5	65	9	91	86	88
194660	A	56	22	4	34	8	82	76	79
861110	A	59	25	4	34	4	84	88	86
26720	AGC	59	26	5	33	5	81	85	83
107750	AGC	25	8	1	17	0	88	100	96
119300	AGC	50	21	4	29	5	81	83	82
136010	AGC	53	21	1	32	1	95	97	96
137550	AGC	16	2	0	14	3	100	79	81
149100	AGC	73	34	5	39	1	85	97	92
153400	AGC	52	22	4	30	1	82	97	90
307100	AGC	90	40	6	50	7	85	86	86
328520	AGC	51	19	0	32	3	100	91	94
365520	AGC	51	22	2	29	2	91	93	92
408760	AGC	74	33	3	41	3	91	93	92
426520	AGC	179	78	5	101	9	94	91	92
426605	AGC	61	27	2	34	4	93	88	90
453920	AGC	39	16	2	23	1	88	96	92
485550	AGC	52	21	4	31	2	81	94	88
130300	AH	94	33	14	61	24	58	61	60
130350	AH	72	26	7	46	18	73	61	65
624515	AH	60	28	11	32	14	61	56	58
624550	AH	59	27	10	32	14	63	56	59
926100	AH	97	42	22	55	19	48	65	58
938210	AH	89	32	17	57	27	47	53	51
35450	G	70	32	7	38	4	78	89	84
65100	G	112	49	12	63	15	76	76	76
82870	G	66	28	2	38	9	93	76	83
83900	G	136	66	3	70	12	95	83	89
190220	G	114	53	15	61	12	72	80	76
263100	G	109	48	11	61	10	77	84	81
269030	G	56	22	3	34	6	86	82	84
376220	G	63	28	5	35	4	82	89	86
376240	G	62	27	6	35	6	78	83	81
449890	G	22	9	3	13	5	67	62	64
941000	G	209	87	12	122	18	86	85	86
6009	H	23	11	2	12	2	82	83	83
57500	H	51	25	5	26	6	80	77	78
183700	H	15	9	0	6	1	100	83	93
190260	H	85	43	6	42	7	86	83	85
520400	H	87	44	1	43	8	98	81	90
546000	H	142	71	9	71	12	87	83	85
629100	H	81	40	4	41	5	90	88	89
883560	H	11	5	1	6	2	80	67	73
Overall mean similar classification (%)							**82**	**82**	**82**

3.4.3 Selection of thirty sampling points for isotope monitoring

The decision tree model performance on the 47 sampling points implied that some sampling points showing low CCI may indicate a low reconstrutability given by physico-chemical data compared to the expert classification. Thus, expert knowledge needs to be retrieved based on an independent method for potentially more accurate NO_3^- classes. Stable N and O isotope methods have been greatly used to identify NO_3^- sources in water. It has been accepted that the $\delta^{15}N$-and $\delta^{18}O$-NO_3^- values of NO_3^- can provide useful information on the sources of NO_3^- in water, as different NO_3^- sources (fertilizer, manure, industrial/septic waste, and atmospheric N deposition) have unique isotopic ratios of N ($^{15}N/^{14}N$) and O ($^{18}O/^{16}O$). Therefore, 30 sampling points have been selected in Table 3-5 for isotope monitoring.

Table 3-5: Selection of 30 sampling points for isotope monitoring from October 2007 to September 2009.

VMM number	Basin	NO_3^- source class	CCI (%)
13000	A	Canals around Ghent	87
13500	A	Canals around Ghent	85
17000	A	Canals around Ghent	78
24050	A	Polders of Bruges	88
194660	A	Lower-Scheldt	79
861110	A	Yzer	86
26720	AGC	Polders of Bruges	83
153400	AGC	Meuse	90
307100	AGC	Nete	86
365520	AGC	Dijle Zenne	92
408760	AGC	Demer	92
426520	AGC	Demer	92
130300	AH	Meuse	60
130350	AH	Meuse	65
624515	AH	Leie	58
624550	AH	Leie	59
926100	AH	Yzer	58
938210	AH	Yzer	51
65100	G	Meuse	76
83900	G	Meuse	89
190220	G	Lower-Scheldt	76
263100	G	Nete	81
376220	G	Dijle Zenne	86
449890	G	Demer	64
6009	H	Polders of Bruges	83
57500	H	Canals around Ghent	78
183700	H	Lower-Scheldt	93
520400	H	Dender	90
629100	H	Leie	89
883560	H	Polders of Bruges	73

The criteria for representative sampling point selection were (a) selecting sampling points in a relatively wide range of CCI (51% to 93%), (b) selecting multiple sampling points per expert class to maximize including sampling points from each basin in the corresponding class, and (c) selecting sampling points distributed over the whole of Flanders in different basins. Since the least number of sampling points is 6 in class AH and to keep uniform in each class , 30 sampling points (5 classes x 6 sampling points per class) in total were selected for isotope and physicochemical monitoring. The sampling period last for 2 years (October 2007 – September 2009). The following stable isotopes and physicochemical parameters have been analyzed during the monitoring period: N isotope ratio ($\delta^{15}N$- NO_3^-), O isotope ratio ($\delta^{18}O$- NO_3^-), boron isotope ratio ($\delta\ ^{11}B$), temperature (T), electrical conductivity at 20 °C (EC 20), pH, dissolved oxygen (DO), oxygen saturation (O_{2sat}) , chloride (Cl^-), ammonium (NH_4^+), nitrite (NO_2^-), nitrate (NO_3^-) and phosphate (PO_4^{3-}).

3.5 Conclusions

The decision tree model developed in this study is useful in extracting key information from a relatively large data set composed of 47 sampling points with 10 physicochemical parameters during a long historical monitoring. The decision tree model provided information on reconstrutability given by physico-chemical data compared to the expert classification. Overall 82% CCI has been obtained. However, some sampling points showed a low CCI. Therefore, we suggested that the expert classification required further retrieved via isotope fingerprinting. Finally, evaluation of the 47 individual sampling points via a decision tree model provided reasonable criteria for selecting representative sampling points for isotope monitoring.

Chapter 4:

Comparison of isotope techniques used to determine N and O isotope ratios of nitrate in surface water

This chapter has been edited from:

Xue, D., De Baets, B., Botte, J., Vermeulen, J., Van Cleemput, O. and Boeckx, P. (2010) Comparison of the silver nitrate and bacterial denitrification methods for the determination of nitrogen and oxygen isotope ratios of nitrate in surface water. Rapid Commun. Mass Spectrom. 24, 833–840.

4.1 Abstract

This chapter we will find a suitable method for measuring N and O isotopes of NO_3^- for isotope monitoring. Both the silver nitrate ($AgNO_3$) method and the bacterial denitrification method are frequently used analytical techniques to determine $\delta^{15}N$- and $\delta^{18}O$-NO_3^- in aqueous samples. The $AgNO_3$ method is applicable for freshwater and requires a concentration of 100-200µmol of NO_3^- for isotope determination. The bacterial denitrification method is applicable for seawater and freshwater and KCl extracts of soils with a NO_3^- concentration as low as 1 µmol. We carried out a thorough method comparison using 42 real surface water samples having a wide range of $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values. Various correction pairs using three international references and blanks were used to correct raw $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values. No significant difference between the corrected data was observed when using various correction pairs for each analytical method. Both methods also showed an excellent repeatability with high intraclass correlation coefficients (ICC). The ICC of the $AgNO_3$ method was 0.992 for $\delta^{15}N$ and 0.970 for $\delta^{18}O$. The ICC of the bacterial denitrification method was 0.995 for $\delta^{15}N$ and 0.954 for $\delta^{18}O$. Moreover, a positive linear relation with a high correlation coefficient ($r \geq 0.88$) between both methods was found for $\delta^{15}N$- and $\delta^{18}O$-NO_3^-. The comparability of both methods was assessed by the Bland-Altman technique using 95% limits of agreement. The average difference between results obtained by the bacterial denitrification and the $AgNO_3$ method for $\delta^{15}N$ was -1.5‰ with 95% limits of agreement -3.6 and +0.5‰. For $\delta^{18}O$ this was +2.0‰, with 95% limits of agreement -3.3 and +7.3‰. We found that for $\delta^{15}N$ and for $\delta^{18}O$, 97% of the differences fell within these 95% limits of agreement. In conclusion, the $AgNO_3$ and the bacterial denitrification methods are highly correlated and statistically interchangeable. However, on average, the bacterial denitrification method yields $\delta^{15}N$ and $\delta^{18}O$ values that deviate -1.5‰ and +2‰ from the the $AgNO_3$ method.

4.2 Introduction

Nitrate (NO_3^-) contamination of water systems is recognized as a global environmental problem. Stable nitrogen (N) and oxygen (O) isotope ratios of NO_3^- are considered as a powerful tool to identify potential NO_3^- pollution sources. Depending on the matrix, concentrations and experimental conditions, several methods are being used for $\delta^{15}N$ and $\delta^{18}O$ analysis in NO_3^-. Silva et al.(2000) presented an ion-exchange method, the so-called $AgNO_3$

method, in which NO_3^- is first loaded on an ion-exchange resin and subsequently eluted, neutralized and purified for analysis of $\delta^{15}N$- and $\delta^{18}O$-NO_3^-. However, this method is relatively labor- and cost-intensive and not suitable for seawater samples and KCl extracts and for freshwater samples with low NO_3^- concentration (Sigman et al., 2001; Xue et al., 2009). Hence, a new technique, the bacterial denitrification method, has been developed. This method allows for simultaneous determination of $\delta^{15}N$ and $\delta^{18}O$ in N_2O produced from the conversion of NO_3^- in water by denitrifying bacteria lacking active N_2O reductase. Although this method may overcome the drawbacks of the $AgNO_3$ method, uncertainties with regard to O isotope fractionation and O exchange between nitrogen oxide intermediates and water during the conversion process require proper correction (Casciotti et al., 2002).

To our knowledge, Sigman et al. (2001) and Casciotti et al.(2002) measured the isotopic composition of NO_3^- in some groundwater and precipitation samples to compare the bacterial denitrification method and combustion-based methods (NO_3^- was extracted and purified using different sample preparation technique). However, they didi not use the commonly used pyrolysis-based $AgNO_3$ method and limited number of real water samples. In this chapter, we analyzed 42 real surface water samples, showing a relatively wide range of $\delta^{15}N$- and $\delta^{18}O$-NO_3^- for comparison of the $AgNO_3$ method and the bacterial denitrification method. Here, the $AgNO_3$ method is considered as the reference method, while the bacterial denitrification method will be evaluated as the alternative method. The objective of this chapter was to statistically elucidate whether results obtained by the bacterial denitrification method are similar to those obtained by the $AgNO_3$ method.

4.3 Material and methods

4.3.1 Sample selection

Surface water from three sampling points, 365520$^{\#}$ (50°45'57" N, 4°12' E), 449890$^{\#}$ (50°55'12" N, 5°12' E) and 938210$^{\#}$ (50°57'59" N, 2°53' E) from the MAP (Manure Action Plan) monitoring network operated by VMM (Flemish Environment Agency) was used. Sampling points 365520#, 449890# and 938210# are predominantly influenced by agriculture, greenhouses in an agricultural and a mixture of agriculture with horticulture nitrate source, respectively. The mean NO_3^- concentrations of the three sampling points during the historical monitoring period (2002 - 2009) were 6.5 ± 2.0 mg NL^{-1} for sampling point 365520#, 44.2 ± 40.0 mg NL^{-1} for sampling point 449890# and 21.3 ± 12.3 mg NL^{-1} for sampling point 938210#. Surface water was sampled monthly from October 2007 to November 2008.

Samples were collected in 1L polyethylene bottles and stored in a freezer before filtration (0.45μm), NO_3-N determination and analysis of $\delta^{15}N$- and $\delta^{18}O$-NO_3^-.

4.3.2 The ion-exchange resin method or $AgNO_3$ method

The $AgNO_3$ method is used for concentration and purification of NO_3^- in surface and groundwater samples for simultaneous determination of $\delta^{15}N$ and $\delta^{18}O$. In brief, first, NO_3^- is concentrated by passing water samples through cation and anion exchange resin columns. Thereafter, NO_3^- is eluted with hydrochloric acid, neutralized with silver oxide and then filtered to remove the AgCl. For accurate $\delta^{18}O$-NO_3^- analysis, all non-nitrate oxygen-bearing anions (e.g. SO_4^{2-}, CO_3^{2-} and PO_4^{3-}) are removed from the sample by adding $BaCl_2$ to the $AgNO_3$ solution and filtration of the precipitate. Subsequently, the filtered solution is passed through a cation exchange resin to remove excess Ba^{2+} ions and re-neutralized with Ag_2O. The resulting solution is freeze-dried to produce $AgNO_3$ salt for isotope analysis. Finally, $\delta^{15}N$ and $\delta^{18}O$ are simultaneously analyzed via a TC/EA (thermal conversion/elemental analyzer) coupled to an isotope ratio mass spectrometer (20-20, SerCon Ltd, UK). The prepared $AgNO_3$ samples are pyrolized at 1400^oC in a molybdenum-lined, aluminum oxide reduction tube filled with glassy carbon and topped with a glassy carbon crucible. Produced N_2 and CO gases are separated via a 1m gas chromatographt (GC) column (E3030, Elemental Microanalysis Ltd, UK) at a temperature of 50^oC, helium (He) carrier pressure 1.6bar, He flow retention time ~250 sec and sample analysis time 1000 sec and analyzed via IRMS for $\delta^{15}N$ and $\delta^{18}O$.

4.3.3 The bacterial denitrification method

The summary of the bacterial denitrification method below takes into account minor modifications compared to the original method. In brief, bacterial cultures (*Pseudomonas aureofaciens*), are grown for 6-10 days in amended tryptic soy broth (TSB), then divided into centrifuge tubes of 40 mL aliquots and centrifuged**.** After centrifugation, the supernatant is decanted and reserved, and 4.2mL of the TSB is pipetted back into the tubes to obtain a ~10-fold concentration of bacteria. These tubes are vortexed to ensure homogenized cultures and then transferred as 2 x 2mL aliquots into 20mL headspace vials. The recommended crimp-seals for the headspace vials are replaced with screw-top Teflon-backed silicone septa to avoid leaks. To ensure anaerobic conditions, the headspace vials are purged with N_2 gas for

3h. Water samples, international references (USGS32, USGS34 and USGS35) and an internal working standard (KNO_3) containing 100 nmol N in $NO_3^-$are then injected into the headspace vials and incubated overnight to allow for conversion of dissolved NO_3^- to N_2O. The original method calls for sample sizes as low as 10-20 nmol N in NO_3^-, but we have adjusted this since 100 nmol N in NO_3^- generates more N_2O (~20 ppm) and reduces standard deviations between replicates.. The following day, 0.1mL of 10N NaOH is injected into the headspace vials to stop bacterial activity and to reduce CO_2 removal during N_2O analysis. Instead of direct measurement from the headspace vials, we perform an off-line extraction of the N_2O into evacuated 12 mL exetainer vials. We transfer 6mL of the N_2O generated in the headspace vials into the evacuated exetainer vials and bring these vials up to 1 atm with He. Simultaneous $\delta^{15}N$ and $\delta^{18}O$ analyses of the N_2O produced has been carried out using a Trace gas-preparation unit (ANCA TGII, SerCon Ltd., Crewe, UK) coupled to an isotope ratio mass spectrometer (IRMS) (20-20, SerCon Ltd, UK). The N_2O sample is injected via an auto-sampler and CO/CO_2 is removed using scrubbers (Schutze Reagent which converts CO to CO_2 followed by Carbosorb Granular 6-12 mesh which removes all CO_2). By cryogenic trapping and focusing, the N_2O is compressed onto a capillary column (CP-PoraPlot Q, 25 m, 0.32 mm i.d., 10 µm df; Varian Inc., Palo Alto, CA, USA) at 35°C and subsequently analyzed by IRMS

4.3.4 Stable nitrogen and oxygen isotope determination

Stable isotope ratios are usually expressed in delta (δ) units and a per mil (‰) notation relative to the respective international standards:

$$\delta_{\text{sample}}(‰) = \left(\frac{R_{sample}}{R_{standard}} - 1\right) \times 1000 \qquad (4\text{-}1)$$

where R_{sample} and $R_{standard}$ are the $^{15}N/^{14}N$ or $^{18}O/^{16}O$ ratio of the sample and standard for $\delta^{15}N$ and $\delta^{18}O$, respectively. Values of $\delta^{15}N$ are reported relative to atmospheric air (AIR) and $\delta^{18}O$ values are reported relative to Vienna Standard Mean Ocean Water 2 (VSMOW 2).

Three international references (USGS32, USGS34 and USGS35) with internationally accepted values (Table 4-1) were used to correct raw $\delta^{15}N$ and $\delta^{18}O$ values determined by the $AgNO_3$ method and the bacterial denitrification method.

Table 4-1: $\delta^{15}N$ and $\delta^{18}O$-NO_3^- values of internationally accepted NO_3^- references.

International reference	$\delta^{15}N$-NO_3^-	Source	$\delta^{18}O$-NO_3^-	Source
USGS32	180.0 ± 1.0	IAEA (2004)	25.7 ± 0.4	IAEA (2004)
USGS34	-1.8 ± 0.2	IAEA (2004)	-27.8 ± 0.4	Brand et al. (2009)
USGS35	2.7 ± 0.2	IAEA (2004)	56.8 ± 0.3	Brand et al. (2009)

A $\delta^{15}N$-blank correction method has been applied in which the $\delta^{15}N$ value of the blank (MilliQ water) is calculated based on accepted and measured $\delta^{15}N$ value of the above international references and measured areas of Mass 44 (*m/z* 44) of N_2O (Rock and Ellbert, 2007):

$$\delta^{15}N_{blank} = \frac{\delta^{15}N_{R,measured} \times I44_{R,measured}}{I44_{blank,measured}} - \frac{\delta^{15}N_{R,accepted} \times (I44_{R,measured} - I44_{blank,measured})}{I44_{blank,measured}} \quad (4\text{-}2)$$

where I_{44} is the measured area of Mass 44; *R* is the international reference USGS 34 or USGS 35. The international reference USGS32 is not appropriate for the $\delta^{15}N$ blank calculation because of rather high $\delta^{15}N$ value (180.0‰).

Then a raw $\delta^{15}N$ value of a water sample can be corrected based on the following equation:

$$\delta^{15}N_{sample,corrected} = \frac{\delta^{15}N_{sample,measured} \times I44_{sample,measured} - \delta^{15}N_{blank} \times I44_{blank,measured}}{I44_{sample,measured} - I44_{blank,measured}} \quad (4\text{-}3)$$

Since O isotope fractionation and O exchange can occur in the bacterial denitrification method, the blank correction method is not fully suitable for $\delta^{18}O$. Therefore, a correction factor method is applied to correct raw $\delta^{18}O$-NO_3^- values of water samples and this method is also suitable for raw $\delta^{15}N$-NO_3^- correction. The correction factor (CF) is computed by the ratio of the difference between the accepted $\delta^{15}N$ or $\delta^{18}O$ values of two international references and the difference between the measured values of the two international references (Casciotti et al., 2002; Rock and Ellbert, 2007)

$$CF_{15N} = \frac{\delta^{15}N_{R1,accepted} - \delta^{15}N_{R2,accepted}}{\delta^{15}N_{R1,measured} - \delta^{15}N_{R2,measured}} \quad (4\text{-}4)$$

where R_1 and R_2 can be a combination of USGS32 - USGS35 and USGS32 – USGS34, which provide relatively wide correction windows for $\delta^{15}N$ correction, and

$$CF_{18O} = \frac{\delta^{18}O_{R1,accepted} - \delta^{18}O_{R2,accepted}}{\delta^{18}O_{R1,measured} - \delta^{18}O_{R2,measured}} \quad (4\text{-}5)$$

where R_1 and R_2 can be a combination of USGS34 – USGS35 and USGS34 – USGS32, which expands the correction windows for $\delta^{18}O$ corrections.

Then a raw $\delta^{15}N$ and $\delta^{18}O$ value of a water sample can be corrected based on the following equation:

$$\delta^{15}N_{sample,corrected} = \delta^{15}N_{R,accepted} + (\delta^{15}N_{sample,measured} - \delta^{15}N_{R,measured}) \times CF_{15N} \quad (4\text{-}6)$$

and

$$\delta^{18}O_{sample,corrected} = \delta^{18}O_{R,accepted} + (\delta^{18}O_{sample,measured} - \delta^{18}O_{R,measured}) \times CF_{18O} \quad (4\text{-}7)$$

where R can be any of the international references used in the corresponding correction factor calculation. Based on Eqs. (4-2) – (4-4) and (4-6), four (f_1 - f_4) corrected $\delta^{15}N$-NO_3^- values can be obtained via a combination of a blank and USGS34 (f_1), via a combination of a blank and USGS35 (f_2), via a combination of USGS32 and USGS35 for CF_{15N} (f_3), and via a combination of USGS32 and USGS34 for CF_{15N} (f_4). Thus, the final corrected $\delta^{15}N$ value of a water sample has been computed based on the average of the four correction pairs available:

$$f_{15N} = \frac{f_1 + f_2 + f_3 + f_4}{4} \quad (4\text{-}8)$$

Based on Eqs. (4-5) and (4-7), two (g_1- g_2) corrected $\delta^{18}O$-NO_3^- values can be obtained via a combination of USGS34 and USGS35 for CF_{18O} (g_1) and via a combination of USGS34 and USGS32 for CF_{18O} (g_2). Thus, the final corrected $\delta^{18}O$ value of a water sample has been computed based on the average of the two correction pairs:

$$g_{18O} = \frac{g_1 + g_2}{2} \quad (4\text{-}9)$$

The correction methods were both applied to the $AgNO_3$ method and the bacterial denitrification method for their comparison.

4.3.5 Statistical analysis

As we used two replicates for each water sample in both methods, an intraclass correlation coefficient (ICC) (Shrout and Fleiss, 1979) could be used to assess the repeatability of both the $AgNO_3$ method and the bacterial denitrification method. ICC is applied in the framework of analysis of variance (ANOVA) and can be calculated following different approaches based on the assumptions of the specific ANOVA model. In our study, the ICC is dependent on the one-way random effects model and was calculated as follows (Kish, 1965; McGraw, 1996):

$$ICC = \frac{MS_{between} - MS_{within}}{MS_{between} + (k-1)MS_{within}} \tag{4-10}$$

where k represents the number of replications and $MS_{between}$ and MS_{within} are the means of the sums of squares from the one-way ANOVA model for between and within replicates of water samples, respectively. ICC values equal to 0 represent no repeatability, while 1 represents perfect repeatability. In accordance with Landis and Koch (1977), the following ICC interpretation scale was used: poor to fair (below 0.4), moderate (0.41 – 0.60), excellent (0.61 – 0.80), and almost perfect (0.81 – 1).

In addition, to check whether there was a significant difference between corrected $\delta^{15}N$ and $\delta^{18}O$ values using various correction pairs of international references and blanks, the Tukey HSD (Tukey Honest Significant Difference) test was applied to perform a multiple comparison.

The Bland-Altman technique (Bland and Altman, 1986) is useful for assessing the agreement between two measurement methods. This technique has been applied here to compare results obtained via the $AgNO_3$ method and the bacterial denitrification method. The average ($\bar{d}$) and standard deviation (s_d) of the difference (d) between measurement results of both methods on the same water samples have been computed. If the differences are normally distributed, and 95% of these differences fall between $\bar{d} - 1.96\, s_d$ and $\bar{d} + 1.96\, s_d$, which represent 95% limits of agreement, we can use the two analytical methods interchangeably.

4.4. Results and discussion

4.4.1 Repeatability of the $AgNO_3$ method and the bacterial denitrification method

Repeatability is relevant to the comparison of methods, because it could affect the magnitude of agreement. When methods have a poor repeatability, this could lead to poor agreement between the two methods. Even if the results obtained by the two methods agree very close on average, poor repeatability of one method would lead to poor agreement between the methods for individuals. The computed intraclass correlation ICC values of the $AgNO_3$ method were 0.992 and 0.970 for $\delta^{15}N$ and $\delta^{18}O$, respectively. The ICC values of the bacterial denitrification method were 0.995 and 0.954 for $\delta^{15}N$ and $\delta^{18}O$, respectively. Thus, the high ICC values (ICC > 0.9) of both methods show an almost perfect repeatability for determination of $\delta^{15}N$- and $\delta^{18}O$-NO_3^- .

4.4.2 Multiple comparison using different international references for correction

Corrected $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of all 42 water samples are summarized in this Table 4-2. There are two sample measurements in this table (365520$^{\#}$: Nov-07 and Mar-08) with outlier $\delta^{18}O$ values, which are not considered in this study. The Tukey HSD test results for corrected $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values using various correction pairs are shown in Figures 4-1 and 4-2. It is obvious that all confidence intervals for differences of the means under various correction pairs contain zero corresponding to Tukey-corrected *p*-values greater than 0.05. This indicates that the differences between different correction pairs were not significant at the 95% level. Thus, all corrections applied give similar results.

Table 4-2: Corrected $\delta^{15}N\text{-}NO_3^-$ and $\delta^{18}O\text{-}NO_3^-$ for all measurements of 42 surface water samples using different correction pairs and applying the $AgNO_3$ and bacterial denitrification method.

Sampling points	Sampling time	$\delta^{15}N$-$AgNO_3$ (‰)					$\delta^{18}O$-$AgNO_3$ (‰)			$\delta^{15}N$-Bacterial (‰)					$\delta^{18}O$-Bacterial (‰)		
		f_1	f_2	f_3	f_4	$f15_N$	g_1	g_2	$g18_O$	f_1	f_2	f_3	f_4	$f15_N$	g_1	g_2	$g18_O$
365520	Oct-07	10.1	9.6	&	&	9.8	6.8	&	6.8	11.2	10.8	11.0	11.3	11.0	7.7	9.4	8.6
		11.9	11.5	&	&	11.7	4.9	&	4.9	11.1	10.7	10.8	11.1	10.9	7.9	9.7	8.8
365520	Nov-07	12.6	12.2	&	&	12.4	4.5	&	4.5	12.3	11.8	11.6	11.8	11.9	-9.4	-9.1	-9.2
		14.1	13.6	&	&	13.8	3.2	&	3.2	12.3	11.8	11.8	11.9	11.9	-9.1	-8.8	-9.0
365520	Dec-07	12.5	10.3	&	&	11.4	3.4	&	3.4	10.5	10.0	9.5	9.7	10.0	6.9	7.4	7.2
		12.1	10.0	&	&	11.1	4.3	&	4.3	10.5	10.0	9.9	10.0	10.1	7.3	7.8	7.6
365520	Jan-08	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
365520	Feb-08	11.6	12.0	11.6	12.0	11.8	9.7	7.0	8.3	9.9	9.2	8.4	9.2	9.2	5.7	6.5	6.1
		11.6	12.1	11.7	12.1	11.8	8.8	6.1	7.4	9.8	9.1	8.5	9.2	9.2	5.1	5.9	5.5
365520	Mar-08	9.1	9.5	9.1	9.5	9.3	6.8	4.3	5.6	7.5	6.8	5.9	6.6	6.7	3.3	4.1	3.7
		9.0	9.4	9.0	9.4	9.2	6.3	3.8	5.0	5.5	4.7	3.9	4.7	4.7	-19.5	-19.3	-19.4
365520	Apr-08	10.5	10.9	10.5	10.9	10.7	9.3	6.6	7.9	9.9	9.2	8.5	9.3	9.3	9.7	10.6	10.1
		11.0	11.5	11.1	11.5	11.3	8.8	6.1	7.5	9.5	8.8	7.8	8.5	8.7	4.7	5.5	5.1
365520	May-08	11.6	12.0	11.6	12.0	11.8	9.1	6.4	7.8	10.2	9.5	8.5	9.3	9.4	4.7	5.5	5.1
		11.9	12.3	12.0	12.4	12.1	9.0	6.4	7.7	10.0	9.3	8.6	9.4	9.3	6.3	7.2	6.7
365520	Jun-08	13.0	13.4	13.0	13.4	13.2	11.0	8.2	9.6	11.4	11.1	11.0	11.1	11.2	6.5	7.3	6.9
		12.8	13.2	12.8	13.2	13.0	10.3	7.6	8.9	11.0	10.6	10.7	10.8	10.8	5.5	6.3	5.9
365520	Jul-08	12.7	12.4	12.7	12.3	12.5	7.6	6.5	7.0	11.3	11.2	11.2	11.0	11.2	13.5	14.0	13.7
		12.7	12.4	12.7	12.3	12.5	6.2	5.2	5.7	11.7	11.5	11.4	11.2	11.4	10.6	11.0	10.8
365520	Aug-08	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
365520	Sep-08	11.7	11.3	11.7	11.3	11.5	7.3	6.2	6.7	11.2	11.1	11.2	11.0	11.1	12.7	13.2	12.9
		12.3	12.0	12.3	11.9	12.1	5.9	4.9	5.4	11.2	11.0	11.2	11.0	11.1	11.0	11.5	11.2
365520	Oct-08	13.3	12.9	13.2	12.8	13.0	10.5	9.4	9.9	12.5	11.4	11.2	12.2	11.8	10.8	12.3	11.5
		13.4	13.0	13.3	12.9	13.1	8.9	7.8	8.3	12.0	10.9	10.5	11.5	11.2	9.8	11.2	10.5
365520	Nov-08	12.1	11.8	12.0	11.6	11.9	6.6	5.5	6.0	12.2	11.1	10.8	11.9	11.5	11.0	12.5	11.7
		12.2	11.8	12.2	11.8	12.0	5.9	4.9	5.4	11.8	10.7	10.5	11.5	11.1	9.6	11.1	10.3
449890	Oct-07	4.8	4.4	&	&	4.6	18.3	&	18.3	4.1	3.7	3.9	4.2	4.0	20.9	23.3	22.1
		4.4	3.9	&	&	4.1	22.4	&	22.4	4.3	3.9	4.0	4.3	4.2	21.2	23.6	22.4
449890	Nov-07	9.0	8.6	&	&	8.8	9.1	&	9.1	7.7	7.1	6.7	6.9	7.1	11.9	12.5	12.2
		8.3	7.9	&	&	8.1	8.2	&	8.2	7.7	7.2	6.5	6.7	7.0	13.9	14.5	14.2

Continued

449890	Dec-07	8.3	6.2	&	&	7.2	10.8	&	10.8	7.2	6.7	6.4	6.5	6.7	13.4	14.0	13.7
		8.6	6.6	&	&	7.6	10.4	&	10.4	6.7	6.1	5.9	6.1	6.2	13.5	14.1	13.8
449890	Jan-08	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
449890	Feb-08	8.4	8.8	8.4	8.8	8.6	18.2	14.9	16.5	6.5	5.8	5.1	5.8	5.8	16.1	17.1	16.6
		8.0	8.4	8.0	8.4	8.2	18.4	15.1	16.8	6.5	5.8	5.1	5.9	5.8	16.6	17.7	17.1
449890	Mar-08	8.8	9.2	8.8	9.2	9.0	13.2	10.2	11.7	3.9	4.3	3.7	2.9	3.7	10.4	11.5	11.0
		8.9	9.3	8.9	9.3	9.1	12.8	9.9	11.3	4.3	4.7	4.2	3.4	4.1	10.4	11.5	10.9
449890	Apr-08	7.4	7.9	7.5	7.9	7.7	18.4	15.1	16.8	4.4	4.2	4.4	4.3	4.3	18.1	15.9	17.0
		7.3	7.7	7.3	7.7	7.5	18.4	15.0	16.7	5.1	5.3	5.0	5.1	5.1	18.7	16.5	17.6
449890	May-08	5.9	6.2	5.9	6.3	6.1	21.7	18.1	19.9	4.4	3.7	3.0	3.8	3.7	19.3	20.4	19.8
		5.4	5.9	5.5	5.9	5.7	23.5	19.7	21.6	4.3	3.6	3.2	3.9	3.8	19.5	20.6	20.1
449890	Jun-08	7.7	8.2	7.8	8.2	8.0	16.7	13.4	15.0	7.9	6.6	6.5	7.9	7.2	14.8	14.7	14.8
		6.9	7.3	6.9	7.3	7.1	18.3	15.0	16.7	7.6	6.2	5.9	7.4	6.8	15.4	15.4	15.4
449890	Jul-08	5.2	5.6	5.2	5.6	5.4	23.7	20.0	21.9	4.7	5.1	4.8	4.6	4.8	23.9	25.5	24.7
		5.3	5.7	5.3	5.7	5.5	24.6	20.8	22.7	5.2	5.5	5.2	5.0	5.2	21.9	23.4	22.7
449890	Aug-08	14.9	14.6	14.9	14.5	14.7	16.8	15.4	16.1	12.2	12.1	11.4	11.1	11.7	19.8	20.3	20.0
		15.0	14.6	14.9	14.6	14.8	18.4	17.0	17.7	12.7	12.5	11.8	11.5	12.1	17.2	17.7	17.5
449890	Sep-08	5.3	4.8	5.3	4.9	5.1	22.6	21.1	21.8	4.9	4.8	4.8	4.5	4.8	30.9	31.6	31.3
		5.7	5.4	5.7	5.3	5.5	20.5	19.0	19.7	4.4	4.3	4.2	4.0	4.2	29.6	30.3	30.0
449890	Oct-08	6.0	5.6	6.0	5.6	5.8	22.2	20.6	21.4	5.1	4.0	4.0	4.9	4.5	25.3	27.4	26.4
		6.0	5.6	5.9	5.5	5.8	22.4	20.9	21.7	4.8	3.6	3.7	4.6	4.2	26.1	28.1	27.1
449890	Nov-08	11.8	11.4	11.7	11.3	11.6	12.2	11.0	11.6	10.7	9.6	9.2	10.3	9.9	12.3	13.9	13.1
		12.2	11.9	12.2	11.8	12.0	11.1	9.9	10.5	10.6	9.5	9.1	10.1	9.8	12.0	13.6	12.8
938210	Oct-07	13.8	13.4	&	&	13.6	6.6	&	6.6	13.6	13.2	13.5	13.8	13.5	8.0	9.8	8.9
		14.2	13.8	&	&	14.0	5.3	&	5.3	14.6	14.0	14.0	14.2	14.2	7.6	8.1	7.9
938210	Nov-07	15.0	14.6	&	&	14.8	8.6	&	8.6	14.1	13.6	13.4	13.5	13.7	9.3	9.9	9.6
		14.5	14.0	&	&	14.3	8.8	&	8.8	14.5	13.9	14.0	14.2	14.1	13.0	13.6	13.3
938210	Dec-07	15.7	13.6	&	&	14.7	6.6	&	6.6	13.7	13.2	12.9	13.1	13.2	7.7	8.3	8.0
		16.1	14.1	&	&	15.1	5.5	&	5.5	13.9	13.3	13.2	13.4	13.5	8.6	9.2	8.9
938210	Jan-08	15.3	13.3	&	&	14.3	6.5	&	6.5	13.4	12.9	12.5	12.7	12.9	7.6	8.2	7.9
		13.8	11.3	&	&	12.5	6.2	&	6.2	12.8	12.2	12.3	12.4	12.4	8.7	9.3	9.0
938210	Feb-08	14.4	14.8	14.4	14.8	14.6	8.7	6.0	7.4	13.1	12.4	11.7	12.4	12.4	6.4	7.3	6.8
		14.5	14.9	14.5	14.9	14.7	9.6	6.9	8.2	12.8	12.1	11.4	12.1	12.1	6.9	7.7	7.3

Continued

938210	Mar-08	14.8	15.2	14.8	15.2	15.0	12.2	9.3	10.8	15.9	12.5	10.7	15.6	13.7	8.6	9.2	8.9
		15.1	15.6	15.2	15.6	15.4	9.7	7.0	8.4	14.5	11.1	8.7	14.1	12.1	7.7	8.3	8.0
938210	Apr-08	15.8	16.2	15.8	16.2	16.0	12.2	9.3	10.7	16.1	12.3	11.7	16.3	14.1	14.4	19.5	17.0
		16.0	16.5	16.1	16.5	16.3	11.6	8.7	10.1	16.4	12.6	11.7	16.6	14.3	10.7	15.4	13.1
938210	May-08	19.5	19.9	19.5	19.9	19.7	14.4	11.3	12.9	17.9	17.6	17.6	17.7	17.7	11.3	12.2	11.7
		19.5	20.0	19.6	20.0	19.8	14.2	11.2	12.7	18.1	17.7	17.7	17.8	17.8	10.4	11.3	10.8
938210	Jun-08	20.4	20.8	20.5	20.8	20.6	14.4	11.3	12.8	19.4	19.1	19.0	19.1	19.1	11.0	11.9	11.4
		20.8	21.2	20.8	21.2	21.0	14.2	11.2	12.7	20.2	20.1	20.0	19.7	20.0	15.5	16.0	15.7
938210	Jul-08	18.5	18.0	18.4	18.1	18.3	11.4	10.2	10.8	17.3	17.1	17.3	17.1	17.2	15.8	16.3	16.1
		17.4	17.1	17.4	17.0	17.2	9.7	8.6	9.1	17.1	17.0	17.1	16.9	17.0	14.3	14.8	14.6
938210	Aug-08	21.8	21.4	21.7	21.3	21.5	9.3	8.2	8.7	21.8	21.7	21.4	21.1	21.5	14.5	15.0	14.8
		22.1	21.8	22.0	21.7	21.9	10.8	9.6	10.2	21.6	21.5	21.3	21.0	21.3	14.6	15.0	14.8
938210	Sep-08	21.8	21.4	21.7	21.3	21.6	11.3	10.1	10.7	20.5	19.5	19.3	20.3	19.9	13.4	15.0	14.2
		21.8	21.4	21.8	21.4	21.6	11.6	10.4	11.0	21.0	19.9	19.3	20.5	20.2	12.8	14.3	13.5
938210	Oct-08	16.2	15.8	16.1	15.7	16.0	9.8	8.6	9.2	15.5	14.5	14.6	15.5	15.1	10.9	12.4	11.6
		16.0	15.7	16.0	15.6	15.8	8.9	7.7	8.3	15.6	14.5	14.7	15.6	15.1	10.9	12.4	11.6
938210	Nov-08	14.6	14.2	14.5	14.2	14.4	9.4	8.3	8.8	14.2	13.1	12.7	13.8	13.4	9.6	11.1	10.4
		14.6	14.2	14.5	14.1	14.3	8.9	7.8	8.4	13.7	12.6	12.1	13.2	12.9	10.1	11.6	10.8

f_1 represents a raw $\delta^{15}N\text{-}NO_3^-$ value corrected by a combination of blank and USGS34; f_2 represents a raw $\delta^{15}N\text{-}NO_3^-$ value corrected by a combination of blank and USGS35; f_3 represents a raw $\delta^{15}N\text{-}NO_3^-$ value corrected by a combination of USGS32 and USGS35; f_4 represents a raw $\delta^{15}N\text{-}NO_3^-$ value corrected by a combination of USGS32 and USGS34; f_{15_N} represents the $\delta^{15}N\text{-}NO_3^-$ average; g_1 represents a raw $\delta^{18}O\text{-}NO_3^-$ value corrected by a combination of USGS34 and USGS35; g_2 represents a raw $\delta^{18}O\text{-}NO_3^-$ value corrected by a combination of USGS34 and USGS32; and g_{18_O} represents the $\delta^{18}O\text{-}NO_3^-$ average, & represents isotope values not corrected by the corresponding correction pairs; - represents no measurements.

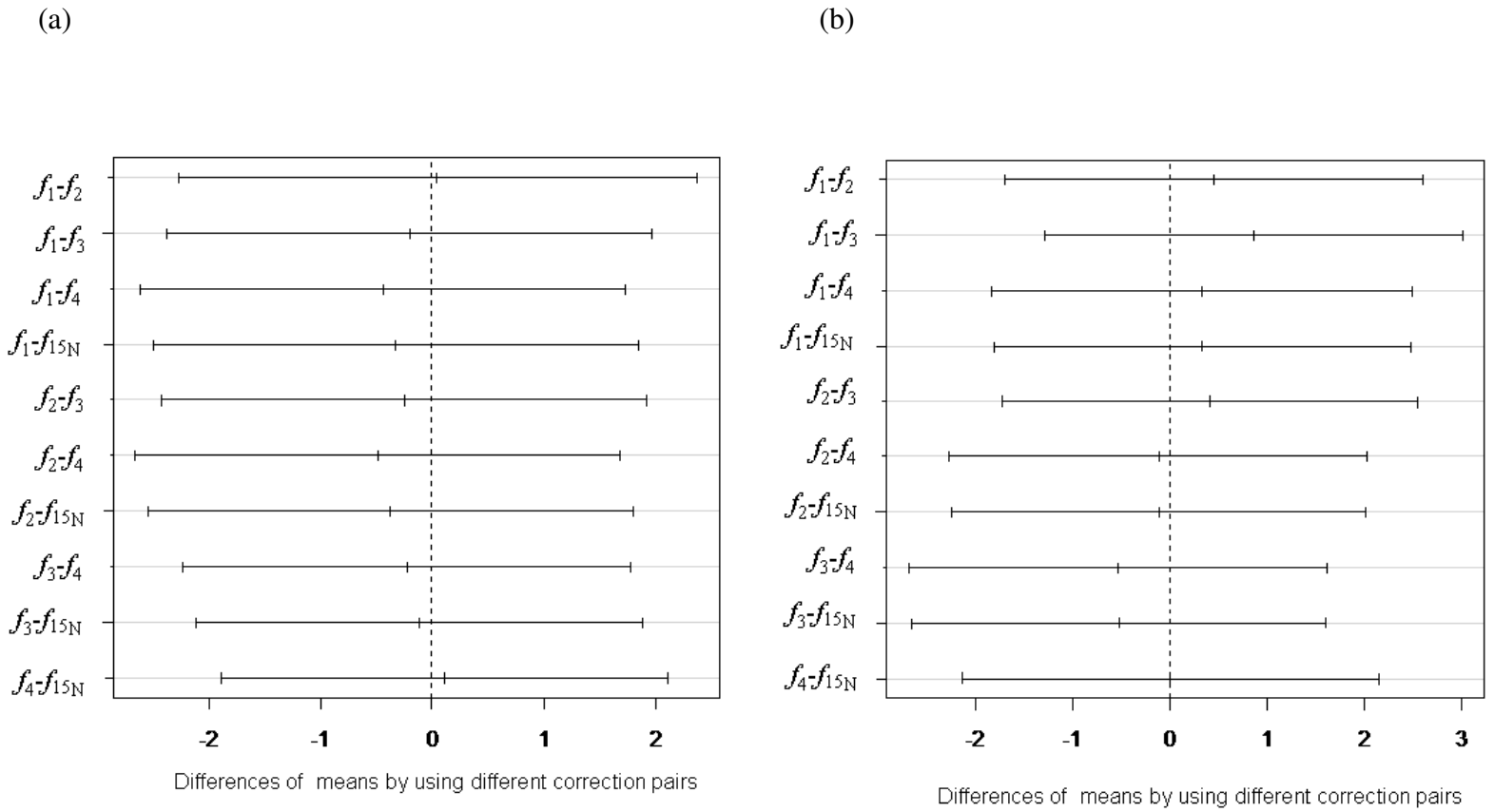


Figure 4-1: 95% confidence intervals on the differences of δ15N-NO_3^- corrected by different correction pairs for the $AgNO_3$ method (a) and the bacterial denitrification method (b). f_1 represents a δ^{15}N-NO_3^- value corrected by a combination of blank and USGS34; f_2 represents a δ^{15}N-NO_3^- value corrected by a combination of blank and USGS35; f_3 represents a δ^{15}N-NO_3^- value corrected by a combination of USGS32 and USGS35; f_4 represents a δ^{15}N-NO_3^- value corrected by a combination of USGS32 and USGS34; f_{15_N} represents the δ^{15}N-NO_3^- average of f_1-f_4

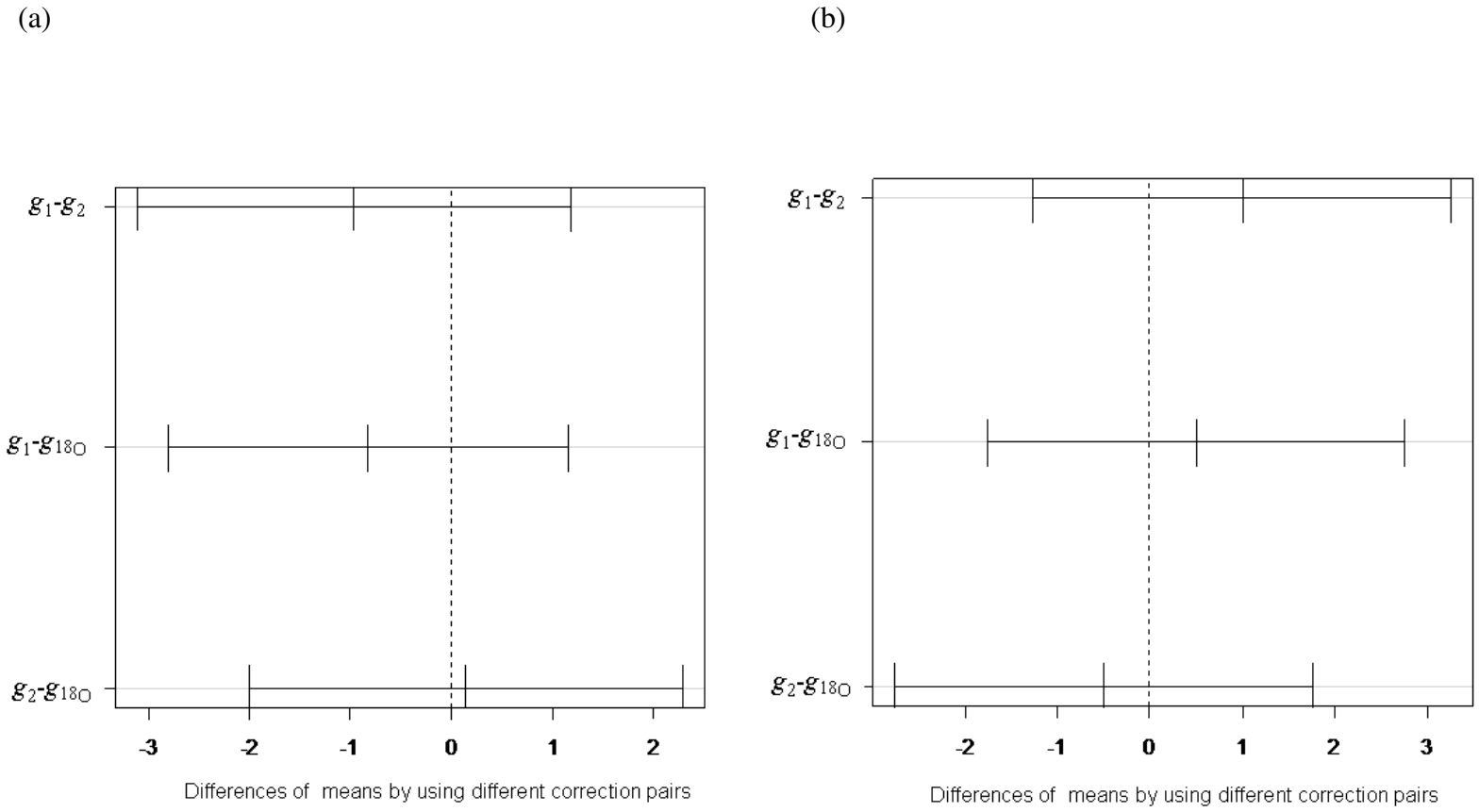
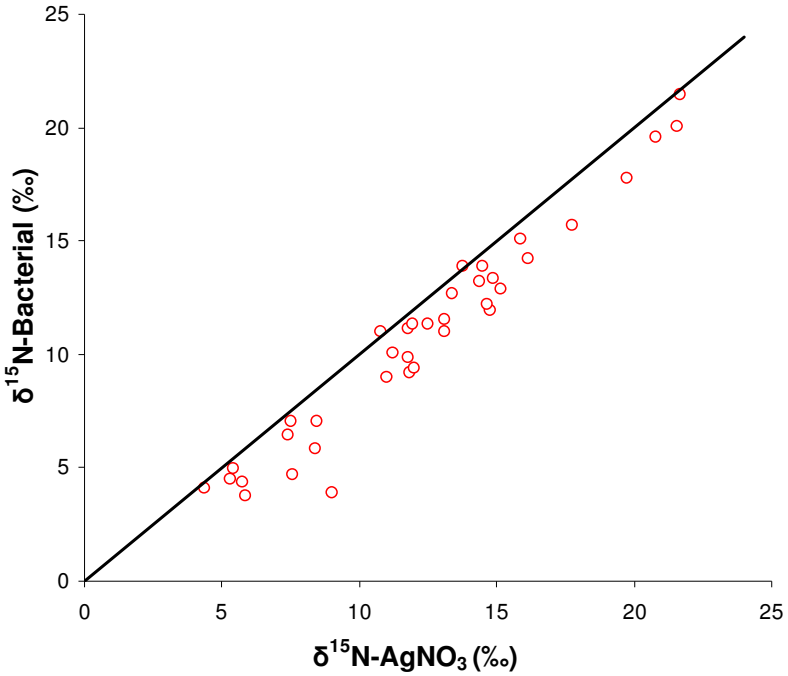


Figure 4-2: 95% confidence intervals on the differences of $\delta^{18}O\text{-}NO_3^-$ corrected by different correction pairs for the $AgNO_3$ method (a) and the bacterial denitrification method (b). g_1 represents a $\delta^{18}O\text{-}NO_3^-$ value corrected by a combination of USGS34 and USGS35; g_2 represents a $\delta^{18}O\text{-}NO_3^-$ value corrected by a combination of USGS34 and USGS32; and g_{18_O} represents the $\delta^{18}O\text{-}NO_3^-$ average of g_1 and g_2.

4.4.3 Comparison of the $AgNO_3$ method and the bacterial denitrification method

The $\delta^{15}N$- and $\delta^{18}O$-NO_3^- of individual water samples determined by the $AgNO_3$ method and the bacterial denitrification method are plotted separately in Figure 4-3. The $\delta^{15}N$-NO_3^- of water samples range from 3.8 to 21.7‰ and $\delta^{18}O$-NO_3^- values are between 3.9 and 30.6‰. Compared with the identity line ($y = x$), it is clear that nearly all $\delta^{15}N$-NO_3^- measurements fall slightly beneath the identity line, which might indicate that $\delta^{15}N$-NO_3^- as determined by the bacterial denitrification method is underestimated compared to the $AgNO_3$ method. On the other hand, the bacterial denitrification method is likely to generate higher $\delta^{18}O$-NO_3^- values than the $AgNO_3$ method, as most of the $\delta^{18}O$-NO_3^- measurements are spread out above the identity line. There is a positive linear relation between both methods for $\delta^{15}N$-NO_3^- ($r = 0.98$, $p < 0.001$) and for $\delta^{18}O$-NO_3^- ($r = 0.88$, $p < 0.001$). However, high correlation coefficients do not necessarily mean a high agreement.

(a)



(b)

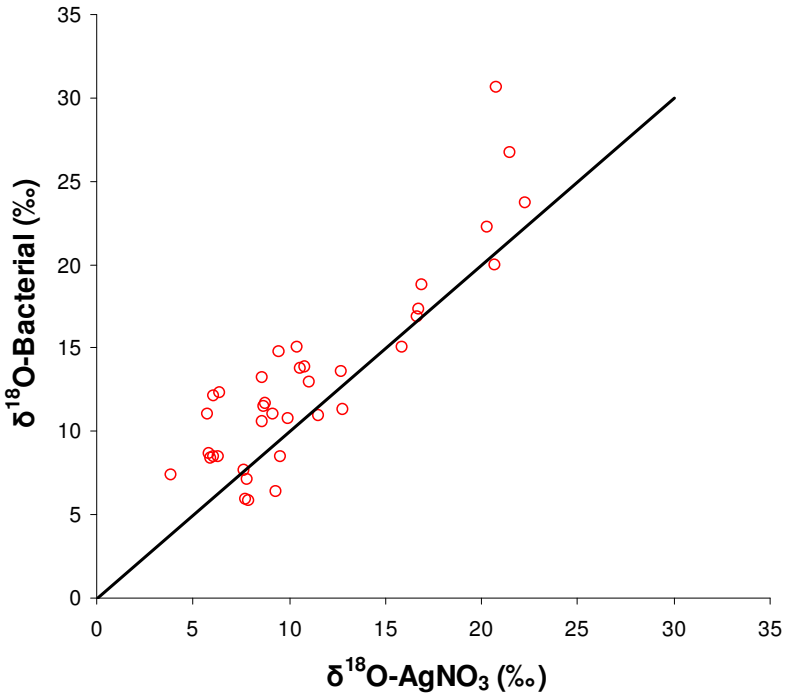
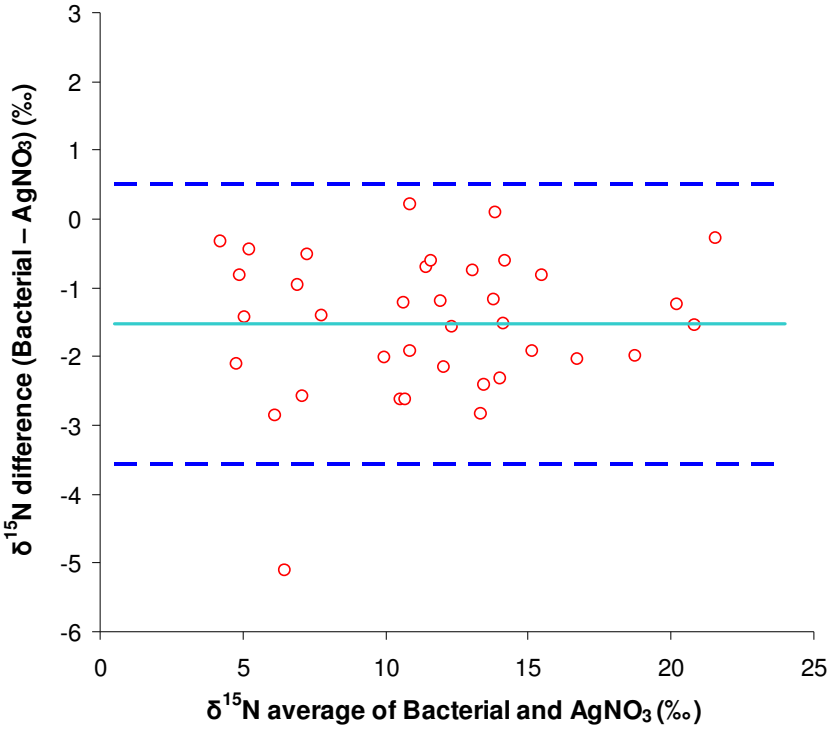


Figure 4-3: Relation between the $AgNO_3$ method and the bacterial denitrification method for $\delta^{15}N\text{-}NO_3^-$ and $\delta^{18}O\text{-}NO_3^-$ determination.

For this purpose, the Bland-Altman technique was applied using 95% limits of agreement (Figure 4-4). The average difference between both measurement techniques for $\delta^{15}N\text{-}NO_3^-$ was -1.5‰, indicating that the bacterial denitrification method might slightly underestimate $\delta^{15}N\text{-}NO_3^-$ compared to the $AgNO_3$ method. The 95% limits of agreement were -3.6 and +0.5‰. The average difference between both methods for $\delta^{18}O\text{-}NO_3^-$ was +2.0‰, indicating that the bacterial denitrification method slightly overestimates $\delta^{18}O\text{-}NO_3^-$ compared to the $AgNO_3$ method. The 95% limits of agreement were -3.3 and +7.3‰. In general, the average difference between the $AgNO_3$ method and the bacterial denitrification method was small and there was no tendency for these differences to vary with the isotope values. The Kolmogorov-Smirnov normality test showed that both $\delta^{15}N\text{-}NO_3^-$ differences (p = 0.87) and $\delta^{18}O\text{-}NO_3^-$ differences (p = 0.69) were normally distributed. The Bland-Altman technique provided limits of agreement within which at least 95% of the differences of the measurements obtained by both methods were expected to fall. We found that for $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ 97% of the differences fell within these 95% limits of agreement, showing that the $AgNO_3$ method and the bacterial denitrification method agreed sufficiently. Thus, both methods are highly correlated and statistically interchangeable. However, on average the bacterial denitrification method tends to underestimate $\delta^{15}N$ (1.5‰) and overestimate $\delta^{18}O$ (2‰) compared to the $AgNO_3$ method. These average deviations for $\delta^{15}N$ and $\delta^{18}O$ are, however, in the same range as the 95% confidence intervals of the differences of the means obtained by using different correction pairs for $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ data correction (Figures 4-1 and 2). This provides further evidence that a 1.5‰ ($\delta^{15}N$) and 2‰ ($\delta^{18}O$) deviation between both methods is reasonable. However, for some individual cases differences can be quite high (> 2‰). This could result from a complex water sample matrix for the three sampling points that were known to be affected by agriculture, greenhouses and a mixture of agricultural and horticultural activities. These three sampling points could be affected by the presence of antibiotics, heavy metals, pesticides, NO_2^- (converted together with NO_3^- by the bacterial denitrification method) etc., all of which could affect the bacterial denitrification method.

The distributions of the differences between both measurement techniques are shown in Figure 4-5. It is clear that the highest frequency of differences for $\delta^{15}N\text{-}NO_3^-$ mainly occurred between -1 and -0.5‰ and the highest frequency of differences for $\delta^{18}O\text{-}NO_3^-$ ranged from 0 to +2‰.

(a)



(b)

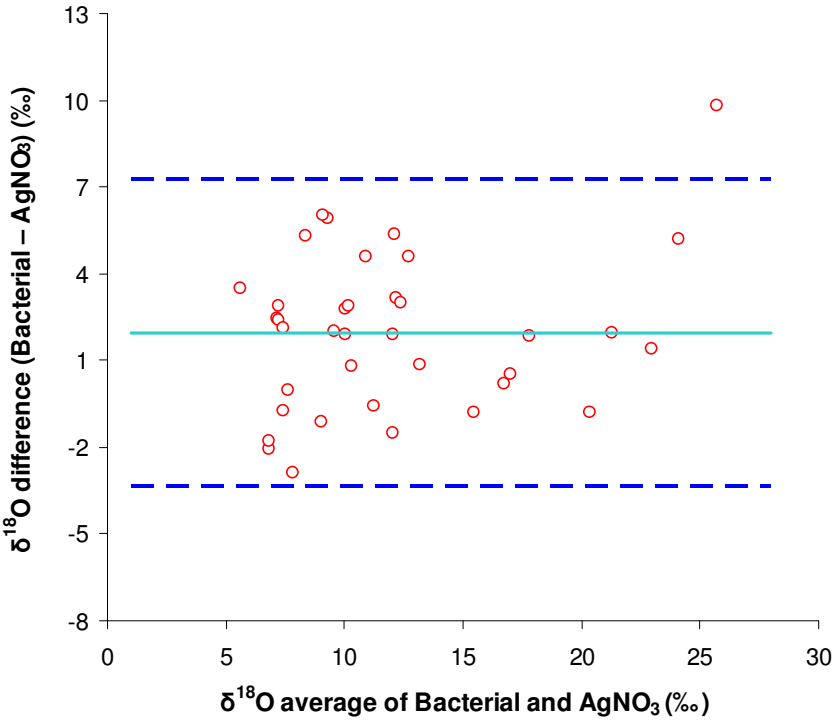
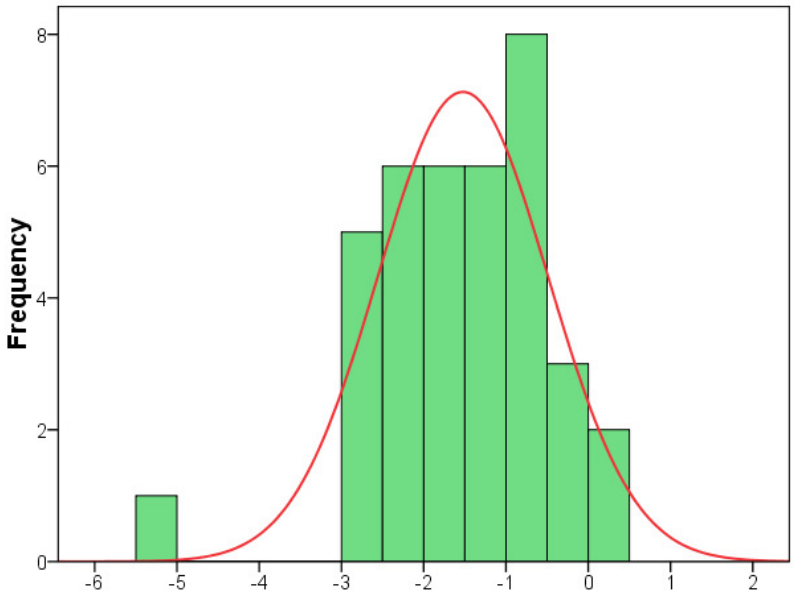


Figure 4-4: Bland-Altman comparison of the $AgNO_3$ method and the bacterial denitrification method for $\delta^{15}N\text{-}NO_3^-$ and $\delta^{18}O\text{-}NO_3^-$ determination. The solid line represents the average difference, while the dashed lines represent the 95% limits of agreement.

a)



(b)

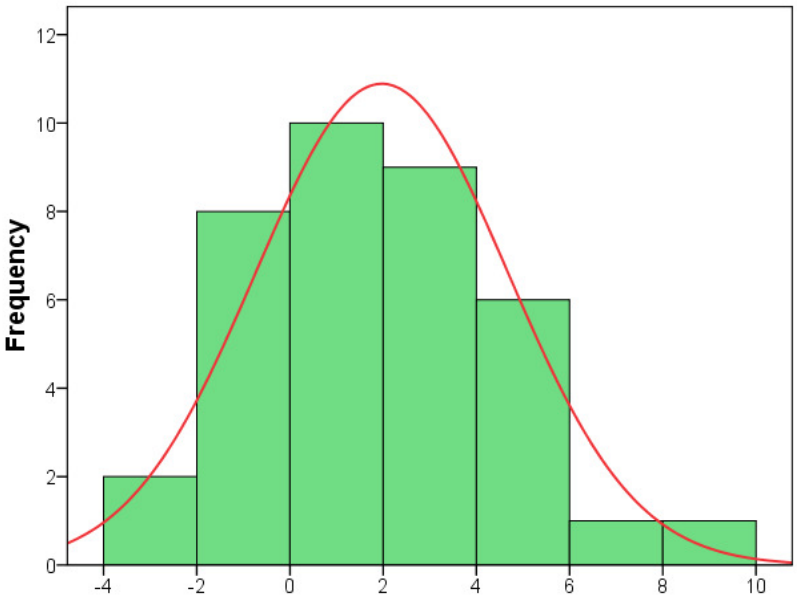


Figure 4-5: Histogram of differences of results obtained by the bacterial denitrification method and the $AgNO_3$ method for $\delta^{15}N\text{-}NO_3^-$ and $\delta^{18}O\text{-}NO_3^-$ determination.

4.5 Conclusions

The $AgNO_3$ method and the bacterial denitrification method are both frequently used analytical techniques for determination of $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ in aqueous samples. We

demonstrated that there is no significant difference in the corrected $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values using various correction pairs for the two methods. ICC analyses showed an excellent repeatability of both methods. Moreover, a positive linear relationship with a high correlation coefficient between both methods has been found for $\delta^{15}N$- and $\delta^{18}O$-NO_3^-. The average measurement difference between the methods indicates that the bacterial denitrification method might underestimate $\delta^{15}N$-NO_3^- and overestimate $\delta^{18}O$-NO_3^- values compared with the $AgNO_3$ method. However, 97% of the differences fall within the 95% limits of agreement. This indicates that the results of the $AgNO_3$ method and those of the bacterial denitrification method are statistically comparable, despite the fact that large differences might occur for some specific samples. The latter case requires further investigation.

Chapter 5:

Error assessment of nitrogen and oxygen isotope ratios of nitrate as determined via the bacterial denitrification method

This chapter has been edited from:

Xue, D., De Baets, B., Vermeulen, J., Botte, J., Van Cleemput, O. and Boeckx, P. (2010) Error assessment of nitrogen and oxygen isotope ratios of nitrate as determined via the bacterial denitrification method. Rapid Commun. Mass Spectrom. 24, 1979–1984.

5.1 Abstract

In chapter 4, we compared the silver nitrate ($AgNO_3$) method and the bacterial denitrification method, which are both frequently used analytical techniques to determine $\delta^{15}N$- and $\delta^{18}O$-NO_3^- in aqueous samples. The $AgNO_3$ method is relatively labor- and cost-intensive and not suitable for seawater samples, KCl extracts and freshwater samples with low NO_3^- concentration (Sigman et al., 2001; Xue et al., 2009). Thus, we will use the bacterial denitrification method to measure water samples. However, proper correction methods with international references (USGS32, USGS34 and USGS35) are needed in this method. As a consequence, it is important to realize that the corrected isotope values are derived from a combination of several other measurements with associated uncertainties. Therefore, it is necessary to consider the propagated uncertainty on the final isotope value. This study demonstrates how to correctly estimate the uncertainty on corrected $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values using a first-order Taylor series approximation. The bacterial denitrification method errors from 33 batches of 561 surface water samples varied from 0.2 to 2.1‰ for $\delta^{15}N$-NO_3^- and from 0.7 to 2.3‰ for $\delta^{18}O$-NO_3^-, which is slightly wider than the machine error, which varied from 0.2 to 0.6‰ for $\delta^{15}N$-N_2O and from 0.4 to 1.0‰ for $\delta^{18}O$-N_2O. The overall uncertainties, which are composed of the machine error and the method error, for the 33 batches ranged from 0.3 to 2.2‰ for $\delta^{15}N$-NO_3^- and from 0.8 to 2.5‰ for $\delta^{18}O$-NO_3^-. In addition, the mean corrected $\delta^{15}N$ and $\delta^{18}O$ values of 132 KNO_3-IWS (internal working standard) measurements were computed as 8.4 ± 1.0‰ and 25.1 ± 2.0‰, which is a slight underestimation for $\delta^{15}N$ and overestimation for $\delta^{18}O$ compared to accepted values ($\delta^{15}N$ = 9.9 ± 0.3‰ and $\delta^{18}O$ = 24.0 ± 0.3‰). The overall uncertainty of the bacterial denitrification method allows the use of this method for source identification of NO_3^-.

5.2 Introduction

Numerous studies (Aravena and Robertson, 1998; Deutsch et al., 2006; Fukada et al., 2004) have shown that stable N and O isotope ratios of NO_3^- ($\delta^{15}N$ and $\delta^{18}O$) can provide useful information on the origin of NO_3^- in water. Thus, precise, accurate, but also cheap and fast analysis of NO_3^- for both $\delta^{15}N$ and $\delta^{18}O$ could offer a tool to improve NO_3^- source identification. The bacterial denitrification method is frequently used for $\delta^{15}N$- and $\delta^{18}O$-NO_3^- analysis (Sigman et al., 2001; Casciotti et al., 2002). This method allows for the simultaneous determination of $\delta^{15}N$ and $\delta^{18}O$ in N_2O that is produced from conversion of NO_3^- in water by

denitrifying bacteria which naturally lack N_2O-reductase activity, and is applicable for seawater, freshwater, groundwater and soil KCl extract samples at natural abundance level. The conversion of NO_3^- into N_2O represents a mass balance reaction for N but not for O. Only 1 of the 6 oxygen atoms present in the initial NO_3^- pool is present in the N_2O produced. Furthermore, exchanges of O between nitrogen oxide intermediates and water during the conversion can bias the O isotopic signal of N_2O (Kool et al., 2007). Thus, correction methods have been developed to correct for potential O isotope fractionation and O exchange during the conversion process (Casciotti et al., 2002; Rock and Ellert, 2007). Final corrected isotope values are derived from a combination of several other measurements of international references and blanks. Each step in the process potentially generates a relatively small amount of error depending on sample preparation, analytical conditions and equipment function. Thus, it is necessary to consider the uncertainty on the final result after correction. Usually the mean and standard deviation from repeated measurements are used to estimate uncertainties. However, in that way, the uncertainty from individual measurements needed for specific corrections and instrument error is not considered, resulting in an underestimation of the uncertainty. Furthermore, a relatively large number of published data only reports uncertainties for standards but not for samples, or does not provide enough detail about the procedure (Jardine and Cunjak, 2005). Thus, the aim of this chapter is to demonstrate how to correctly compute uncertainties on corrected $\delta^{15}N$- and $\delta^{18}O$-NO_3^- measurements in the bacterial denitrification method.

5.3 Material and methods

5.3.1 Method set-up

The bacterial denitrification method allows for simultaneous determination of $\delta^{15}N$ and $\delta^{18}O$ of N_2O produced from the conversion of NO_3^- in water by denitrifying bacteria (e.g. *Pseudomonas aureofaciens)*, which naturally lack active N_2O reductase. In brief, bacterial cultures are grown for 6-10 days in amended tryptic soy broth (TSB), then divided into centrifuge tubes of 40 mL aliquots and centrifuged. After centrifugation, the supernatant is decanted and reserved and 4.2 mL of the TSB is pipetted back into the tubes to obtain a 10-fold concentration of bacteria. These tubes are then vortexed to ensure homogenized cultures and then transferred as 2 x 2 mL aliquots into 20 mL headspace vials. The vials are sealed with screw-top Teflon-backed silicone septa. To ensure anaerobic conditions, the headspace vials are purged with N_2 gas for 3 hours. Water samples, international references (USGS32,

USGS34 and USGS35) and an internal working standard (KNO_3) containing 100 nmol N in NO_3^- are then injected into the headspace vials and incubated overnight to allow complete conversion of NO_3^- to N_2O. The next day, 0.1 mL of 10N NaOH is injected into the headspace vials to stop bacterial activity and to scrub any CO_2 gas in the vial which could interfere with the N_2O measurement. Simultaneous $\delta^{15}N$ and $\delta^{18}O$ analyses of the N_2O produced has been carried out using a Trace gas-preparation unit (ANCA TGII, SerCon Ltd., Crewe, UK) coupled to an isotope ratio mass spectrometer (IRMS) (20-20, SerCon Ltd, UK). The N_2O sample is injected via an auto-sampler and CO/CO_2 is removed using scrubbers (Schutze Reagent which converts CO to CO_2 followed by Carbosorb Granular 6-12 mesh which removes all CO_2). By cryogenic trapping and focusing, the N_2O is compressed onto a capillary column (CP-PoraPlot Q, 25 m, 0.32 mm i.d., 10 µm df; Varian Inc., Palo Alto, CA, USA) at 35°C and subsequently analyzed by IRMS

5.3.2 Stable nitrogen and oxygen isotope determination

Stable isotope ratios are expressed in delta (δ) units and a per mil (‰) notation relative to the respective international standards:

$$\delta_{\text{sample}}(‰) = \left(\frac{R_{sample}}{R_{standard}} - 1\right) \times 1000 \quad (5\text{-}1)$$

where R_{sample} and $R_{standard}$ are the $^{15}N/^{14}N$ or $^{18}O/^{16}O$ ratios of the sample and standard for $\delta^{15}N$ and $\delta^{18}O$, respectively. Values of $\delta^{15}N$ are reported relative to atmospheric air (AIR) and $\delta^{18}O$ values are reported relative to Vienna Standard Mean Ocean Water 2 (VSMOW 2).

Three international references (USGS32, USGS34 and USGS35) are used to correct the raw $\delta^{15}N$ and $\delta^{18}O$ values determined by the bacterial denitrification method. A $\delta^{15}N$-blank correction formula can be used in which the $\delta^{15}N$ value of the blank is calculated based on the accepted and measured $\delta^{15}N$ values of the international references and measured areas of mass 44 of N_2O (Rock and Ellert, 2007):

$$\delta^{15}N_{blank} = \frac{\delta^{15}N_{R,measured} \times I44_{R,measured}}{I44_{blank,measured}} - \frac{\delta^{15}N_{R,accepted} \times (I44_{R,measured} - I44_{blank,measured})}{I44_{blank,measured}} \quad (5\text{-}2)$$

where I_{44} is the measured area of mass 44; *R* is the international reference USGS34 or USGS35. The international reference USGS32 is not appropriate for the $\delta^{15}N$ blank calculation because of the high $\delta^{15}N$ value (180.0‰).

A raw $\delta^{15}N$ value of a water sample can then be corrected based on the $\delta^{15}N_{blank}$:

$$\delta^{15}N_{sample,corrected} = \frac{\delta^{15}N_{sample,measured} \times I44_{sample,measured} - \delta^{15}N_{blank} \times I44_{blank,measured}}{I44_{sample,measured} - I44_{blank,measured}} \quad (5\text{-}3)$$

Since O isotope fractionation and O exchange can occur in the bacterial denitrification method, a blank correction method alone is not suitable for $\delta^{18}O$. Therefore, a correction factor method is applied to correct raw $\delta^{18}O$-NO_3^- values. This method is also suitable for raw $\delta^{15}N$-NO_3^- correction. The correction factor (CF) is estimated according to the ratio of the difference between the accepted $\delta^{15}N$ or $\delta^{18}O$ values of the two international references and the difference between the measured values of the two international references (Casciotti et al., 2002; Rock and Ellert, 2007):

$$CF_{15_N} = \frac{\delta^{15}N_{R1,accepted} - \delta^{15}N_{R2,accepted}}{\delta^{15}N_{R1,measured} - \delta^{15}N_{R2,measured}} \quad (5\text{-}4)$$

where R_1 and R_2 can be the combination of USGS32 – USGS35 and USGS32 – USGS34, which provide relatively wide correction windows for $\delta^{15}N$ correction, and

$$CF_{18_O} = \frac{\delta^{18}O_{R1,accepted} - \delta^{18}O_{R2,accepted}}{\delta^{18}O_{R1,measured} - \delta^{18}O_{R2,measured}} \quad (5\text{-}5)$$

where R_1 and R_2 can be the combination of USGS34 – USGS35 and USGS34 – USGS32, which expands the correction windows for $\delta^{18}O$ corrections.

A raw $\delta^{15}N$ and $\delta^{18}O$ value of a water sample can be corrected based on the following equations:

$$\delta^{15}N_{sample,corrected} = \delta^{15}N_{R,accepted} + (\delta^{15}N_{sample,measured} - \delta^{15}N_{R,measured}) \times CF_{15_N} \quad (5\text{-}6)$$

and

$$\delta^{18}O_{sample,corrected} = \delta^{18}O_{R,accepted} + (\delta^{18}O_{sample,measured} - \delta^{18}O_{R,measured}) \times CF_{18_O} \quad (5\text{-}7)$$

where *R* can be any of the international references used in the corresponding correction factor calculation.

5.3.3 Machine error determination

A machine error is present for all measurements on the TGII IRMS system. Thus, it is essential to consider it in the uncertainty calculation process. The machine error is computed by a series of injections of N_2O-AL (N_2O air liquid) gas with known $\delta^{15}N$ (2.8‰) and $\delta^{18}O$ (35.7‰) values in each individual batch. The standard deviation of the differences between the known and the measured values of N_2O-AL gas is considered as the machine error.

5.3.4 Error propagation

For the bacterial denitrification method, corrected isotope values are derived from measurements of several (*n*) uncorrelated variables. To identify the uncertainty on the final measurement, it is necessary to estimate the uncertainties on these *n* variables and then to determine how these uncertainties propagate through the calculations to produce an uncertainty on the final measurement.

Assuming a variable $t = f(x_1, x_2, \ldots, x_n)$ is a function of *n* uncorrelated variables $x_1, x_2, \ldots, x_n$ with given standard deviations $\sigma_{x_1}, \sigma_{x_2}, \ldots, \sigma_{x_n}$, the uncertainty σ_t of the variable is derived from combining the individual contributions of other variables using the following approximation formula (Mikhail, 1976):

$$\sigma_t \approx \sqrt{\sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i}\sigma_{x_i}\right)^2} \quad (5\text{-}8)$$

which in essence results from a first order Taylor series approximation.

5.3.5 Error propagation in the bacterial denitrification method

There is a mass balance reaction for N in the conversion of NO_3^- to N_2O, so a blank correction method (Eqs. (5-2) and (5-3)) can be applied. A raw $\delta^{15}N\text{-}NO_3^-$ value can be corrected as:

$$f_1 = \delta^{15}N_{sample,corrected,USGS34} = \frac{K \times J - E \times H + B \times H - I \times B}{J - I} \quad (5\text{-}9)$$

and

$$f_2 = \delta^{15}N_{sample,corrected,USGS35} = \frac{K \times J - D \times G + A \times G - I \times A}{J - I} \quad (5\text{-}10)$$

where the variable names are explained in Table 5-1.

Table 5-1: Variables and values for $\delta^{15}N$ and $\delta^{18}O$ correction and error calculation.

Variable	Name	Value	σ
A (‰)	$\delta^{15}N_{USGS35,accepted}$	2.7[&]	0.2
B (‰)	$\delta^{15}N_{USGS34,accepted}$	-1.8[&]	0.2
C (‰)	$\delta^{15}N_{USGS32,accepted}$	180[&]	1.0
D (‰)	$\delta^{15}N_{USGS35,measured}$	6.1	0.3
E (‰)	$\delta^{15}N_{USGS34,measured}$	2.4	0.3
F (‰)	$\delta^{15}N_{USGS32,measured}$	177.3	0.3
G	$I_{44USGS35,measured}$	1.2E-08	2.8E-10
H	$I_{44USGS34,measured}$	1.1E-08	2.8E-10
I	$I_{44blank,measured}$	6.2E-10	2.8E-10
J	$I_{44sample,measured}$	#	2.8E-10
K (‰)	$\delta^{15}N_{sample,measured}$	#	0.3
L (‰)	$\delta^{18}O_{USGS35,accepted}$	56.8[*]	0.3
M (‰)	$\delta^{18}O_{USGS34,accepted}$	-27.8[*]	0.4
N (‰)	$\delta^{18}O_{USGS32,accepted}$	25.7[&]	0.4
O (‰)	$\delta^{18}O_{USGS35,measured}$	87.1	0.8
P (‰)	$\delta^{18}O_{USGS34,measured}$	7.3	0.8
Q (‰)	$\delta^{18}O_{USGS32,measured}$	57.0	0.8
R (‰)	$\delta^{18}O_{sample,measured}$	#	0.8

represents corresponding individual water sample measurements in the example batch

*from Brand et al.

[&]from IAEA

The uncertainties σ_{f1} and σ_{f2} on $\delta^{15}N_{sample,corrected}$ can be computed based on Eq. (5-8):

$$\sigma_{f1}^2 = (\sigma_B \times \frac{\partial f_1}{\partial B})^2 + (\sigma_E \times \frac{\partial f_1}{\partial E})^2 + (\sigma_H \times \frac{\partial f_1}{\partial H})^2 + (\sigma_I \times \frac{\partial f_1}{\partial I})^2 + (\sigma_J \times \frac{\partial f_1}{\partial J})^2 + (\sigma_K \times \frac{\partial f_1}{\partial K})^2 \quad (5\text{-}11)$$

and

$$\sigma_{f2}^2 = (\sigma_A \times \frac{\partial f_2}{\partial A})^2 + (\sigma_D \times \frac{\partial f_2}{\partial D})^2 + (\sigma_G \times \frac{\partial f_2}{\partial G})^2 + (\sigma_I \times \frac{\partial f_2}{\partial I})^2 + (\sigma_J \times \frac{\partial f_2}{\partial J})^2 + (\sigma_K \times \frac{\partial f_2}{\partial K})^2 \quad (5\text{-}12)$$

When combining Eqs. (5-4) and (5-6), a raw $\delta^{15}N$ value of a water sample can also be corrected based on the correction factor method:

$$f_3 = \delta^{15}N_{sample,corrected,USGS32\text{-}USGS35} = \frac{A \times F + K \times C - D \times C - A \times K}{F - D} \quad (5\text{-}13)$$

and

$$f_4 = \delta^{15}N_{sample,corrected,USGS32\text{-}USGS34} = \frac{B \times F + K \times C - E \times C - B \times K}{F - E} \quad (5\text{-}14)$$

where the variable names are explained in Table 5-1.

The uncertainties σ_{f3} and σ_{f4} on $\delta^{15}N_{sample,corrected}$ can be computed as follows:

$$\sigma_{f3}^2 = (\sigma_A \times \frac{\partial f_3}{\partial A})^2 + (\sigma_C \times \frac{\partial f_3}{\partial C})^2 + (\sigma_D \times \frac{\partial f_3}{\partial D})^2 + (\sigma_F \times \frac{\partial f_3}{\partial F})^2 + (\sigma_K \times \frac{\partial f_3}{\partial K})^2 \quad (5\text{-}15)$$

and

$$\sigma_{f4}^2 = (\sigma_B \times \frac{\partial f_4}{\partial B})^2 + (\sigma_C \times \frac{\partial f_4}{\partial C})^2 + (\sigma_E \times \frac{\partial f_4}{\partial E})^2 + (\sigma_F \times \frac{\partial f_3}{\partial F})^2 + (\sigma_K \times \frac{\partial f_4}{\partial K})^2 \quad (5\text{-}16)$$

The final corrected $\delta^{15}N$ value of a water sample is ideally computed using the average of four corrections available:

$$f_{15\mathrm{N}} = \frac{f_1 + f_2 + f_3 + f_4}{4} \tag{5-17}$$

and the overall uncertainty $\sigma_{f_{15\mathrm{N}}}$ can be expressed as:

$$\sigma^2_{f_{15\mathrm{N}}} = (\sigma_A \times \frac{\partial f_{15\mathrm{N}}}{\partial A})^2 + (\sigma_B \times \frac{\partial f_{15\mathrm{N}}}{\partial B})^2 + (\sigma_C \times \frac{\partial f_{15\mathrm{N}}}{\partial C})^2 + (\sigma_D \times \frac{\partial f_{15\mathrm{N}}}{\partial D})^2 + (\sigma_E \times \frac{\partial f_{15\mathrm{N}}}{\partial E})^2 +$$
$$(\sigma_F \times \frac{\partial f_{15\mathrm{N}}}{\partial F})^2 + (\sigma_G \times \frac{\partial f_{15\mathrm{N}}}{\partial G})^2 + (\sigma_H \times \frac{\partial f_{15\mathrm{N}}}{\partial H})^2 + (\sigma_I \times \frac{\partial f_{15\mathrm{N}}}{\partial I})^2 + (\sigma_J \times \frac{\partial f_{15\mathrm{N}}}{\partial J})^2 + (\sigma_K \times \frac{\partial f_{15\mathrm{N}}}{\partial K})^2 \tag{5-18}$$

A raw $\delta^{18}O$ value of a water sample can be corrected based on the correction factor method (Eqs. (5-5) and (5-7)):

$$g_1 = \delta^{18}\mathrm{O}_{sample,corrected,USGS34-USGS35} = \frac{M \times O + R \times L - P \times L - M \times R}{O - P} \tag{5-19}$$

and

$$g_2 = \delta^{18}\mathrm{O}_{sample,corrected,USGS34-USGS32} = \frac{M \times Q + R \times N - P \times N - M \times R}{Q - P} \tag{5-20}$$

where the variable names are explained in Table 5-1.

The uncertainties σ_{g1} and σ_{g2} on $\delta^{18}\mathrm{O}_{sample,corrected}$ can be computed by Eqs. (5-21) and (5-22):

$$\sigma^2_{g1} = (\sigma_L \times \frac{\partial g_1}{\partial L})^2 + (\sigma_M \times \frac{\partial g_1}{\partial M})^2 + (\sigma_O \times \frac{\partial g_1}{\partial O})^2 + (\sigma_P \times \frac{\partial g_1}{\partial P})^2 + (\sigma_R \times \frac{\partial g_1}{\partial R})^2 \tag{5-21}$$

and

$$\sigma^2_{g2} = (\sigma_M \times \frac{\partial g_2}{\partial M})^2 + (\sigma_N \times \frac{\partial g_2}{\partial N})^2 + (\sigma_P \times \frac{\partial g_2}{\partial P})^2 + (\sigma_Q \times \frac{\partial g_2}{\partial Q})^2 + (\sigma_R \times \frac{\partial g_2}{\partial R})^2 \tag{5-22}$$

Then the final corrected $\delta^{18}O$ value of a water sample is ideally computed from two corrections:

$$g_{18_O} = \frac{g_1 + g_2}{2} \tag{5-23}$$

and the overall uncertainty σ_{g18_O} can be expressed as:

$$\sigma^2_{g18_O} = (\sigma_L \times \frac{\partial g_{18_O}}{\partial L})^2 + (\sigma_M \times \frac{\partial g_{18_O}}{\partial M})^2 + (\sigma_N \times \frac{\partial g_{18_O}}{\partial N})^2 + (\sigma_O \times \frac{\partial g_{18_O}}{\partial O})^2 + (\sigma_P \times \frac{\partial g_{18_O}}{\partial P})^2$$
$$+ (\sigma_Q \times \frac{\partial g_{18_O}}{\partial Q})^2 + (\sigma_R \times \frac{\partial g_{18_O}}{\partial R})^2 \tag{5-24}$$

The explicit expressions of the partial derivatives in the above equations are shown in Table 5-2 and Table 5-3 for $\delta^{15}N\text{-}NO_3^-$ and Table 5-4 for $\delta^{18}O\text{-}NO_3^-$.

To demonstrate the error propagation in the bacterial denitrification method, we used data of one batch of seventeen surface water samples. These water samples were collected from different sampling points in the MAP. Furthermore, these water samples are predominantly influenced by different nitrate sources such as agriculture, greenhouses and a mixture of agriculture with horticulture. Samples were collected in 1L polyethylene bottles and stored in a freezer before filtration (0.45µm), NO_3-N determination and measurement of $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$. Three international references USGS32, USGS34 and USGS35 using the most recently reported values were applied for blank correction and calculating correction factors. An internal working standard (KNO_3) was used to check the calculated correction factors. The $\delta^{15}N$ and $\delta^{18}O$ values of KNO_3-IWS (internal working standard) are 9.9 ± 0.3‰ for $\delta^{15}N$ and 24.0 ± 0.3‰ for $\delta^{18}O$ based on 30 KNO_3 measurements. The isotopic composition of KNO_3-IWS was determined by TC/EA-IRMS (thermal conversion/elemental analyzer-isotope ratio mass spectrometry) using IAEA-N3 (4.7 ± 0.2‰ for $\delta^{15}N$ and 25.3 ± 0.3‰ for $\delta^{18}O$) as a reference.

Table 5-2: Partial derivatives in the blank correction method for $\delta^{15}N\text{-}NO_3^-$.

Variable	$f_1 = \frac{K \times J - E \times H + B \times H - I \times B}{J - I}$	$f_2 = \frac{K \times J - D \times G + A \times G - I \times A}{J - I}$
A	—	$\frac{\partial f_2}{\partial A} = \frac{G - I}{J - I}$
B	$\frac{\partial f_1}{\partial B} = \frac{H - I}{J - I}$	—
D	—	$\frac{\partial f_2}{\partial D} = \frac{-G}{J - I}$
E	$\frac{\partial f_1}{\partial E} = \frac{-H}{J - I}$	—
G	—	$\frac{\partial f_2}{\partial G} = \frac{-D + A}{J - I}$
H	$\frac{\partial f_1}{\partial H} = \frac{-E + B}{J - I}$	—
I	$\frac{\partial f_1}{\partial I} = \frac{-B \times (J - I) + (K \times J - E \times H + B \times H - I \times B)}{(J - I)^2}$	$\frac{\partial f_2}{\partial I} = \frac{-A \times (J - I) + (K \times J - D \times G + A \times G - I \times A)}{(J - I)^2}$
J	$\frac{\partial f_1}{\partial J} = \frac{K \times (J - I) - (K \times J - E \times H + B \times H - I \times B)}{(J - I)^2}$	$\frac{\partial f_2}{\partial J} = \frac{K \times (J - I) - (K \times J - D \times G + A \times G - I \times A)}{(J - I)^2}$
K	$\frac{\partial f_1}{\partial K} = \frac{J}{J - I}$	$\frac{\partial f_2}{\partial K} = \frac{J}{J - I}$

Table 5-3: Partial derivatives in the correction factor method for $\delta^{15}N\text{-}NO_3^-$.

Variable	$f_3 = \frac{A\times F + K\times C - D\times C - A\times K}{F - D}$	$f_4 = \frac{B\times F + K\times C - E\times C - B\times K}{F - E}$
A	$\frac{\partial f_3}{\partial A} = \frac{F - K}{F - D}$	—
B	—	$\frac{\partial f_4}{\partial B} = \frac{F - K}{F - E}$
C	$\frac{\partial f_3}{\partial C} = \frac{K - D}{F - D}$	$\frac{\partial f_4}{\partial C} = \frac{K - E}{F - E}$
D	$\frac{\partial f_3}{\partial D} = \frac{-C\times(F - D) + (A\times F + K\times C - D\times C - A\times K)}{(F - D)^2}$	—
E	—	$\frac{\partial f_4}{\partial E} = \frac{-C\times(F - E) + (B\times F + K\times C - E\times C - B\times K)}{(F - E)^2}$
F	$\frac{\partial f_3}{\partial F} = \frac{A\times(F - D) - (A\times F + K\times C - D\times C - A\times K)}{(F - D)^2}$	$\frac{\partial f_4}{\partial F} = \frac{B\times(F - E) - (B\times F + K\times C - E\times C - B\times K)}{(F - E)^2}$
K	$\frac{\partial f_3}{\partial K} = \frac{C - A}{F - D}$	$\frac{\partial f_4}{\partial K} = \frac{C - B}{F - E}$

Table 5-4: Partial derivatives in the correction factor method for $\delta^{18}O\text{-}NO_3^-$.

Variable	$g_1 = \frac{M \times O + R \times L - P \times L - M \times R}{O - P}$	$g_2 = \frac{M \times Q + R \times N - P \times N - M \times R}{Q - P}$
L	$\frac{\partial g_1}{\partial L} = \frac{R - P}{O - P}$	—
M	$\frac{\partial g_1}{\partial M} = \frac{O - R}{O - P}$	$\frac{\partial g_2}{\partial M} = \frac{Q - R}{Q - P}$
N	—	$\frac{\partial g_2}{\partial N} = \frac{R - P}{Q - P}$
O	$\frac{\partial g_1}{\partial O} = \frac{M \times (O - P) - (M \times O + R \times L - P \times L - M \times R)}{(O - P)^2}$	—
P	$\frac{\partial g_1}{\partial P} = \frac{-L \times (O - P) + (M \times O + R \times L - P \times L - M \times R)}{(O - P)^2}$	$\frac{\partial g_2}{\partial P} = \frac{-N \times (Q - P) + (M \times Q + R \times N - P \times N - M \times R)}{(Q - P)^2}$
Q	—	$\frac{\partial g_2}{\partial Q} = \frac{M \times (Q - P) - (M \times Q + R \times N - P \times N - M \times R)}{(Q - P)^2}$
R	$\frac{\partial g_1}{\partial R} = \frac{L - M}{O - P}$	$\frac{\partial g_2}{\partial R} = \frac{N - M}{Q - P}$

5.4 Results and discussion

Table 5-1 shows the variables and values used for raw $\delta^{15}N$- and $\delta^{18}O$-NO_3^- correction and error calculation in the example batch. The isotopic compositions of the three international references, USGS32, USGS34 and USGS35, are internationally accepted and fixed throughout the whole batch, including the corresponding reported uncertainty σ in Table 5-1. The measured $\delta^{15}N$ and $\delta^{18}O$ values of the international references, via the bacterial denitrification method, could vary from batch to batch and may deviate from the accepted values. This is due to the fact that bacteria cultures grown for different sample batches tend to behave slightly different, and potential O isotope fractionation and O exchange during the conversion process contributes to the deviation from the accepted values. Furthermore, all measurements have a machine error and could vary from batch to batch. In Table 5-1, the machine error for $\delta^{15}N$ and $\delta^{18}O$ was 0.3‰ and 0.8‰, respectively. The uncertainty on the measured area of mass 44 of N_2O is 2.8E-10, which is the standard deviation of $I_{44_{blank,measured}}$ in 33 batches (the total number of batches we have analyzed so far). Finally, the raw $\delta^{15}N$ and $\delta^{18}O$ values of the example data have been corrected using the different correction methods. The results and overall uncertainties are shown in Table 5-5. It is clear that the isotope values of the replicates (a, b) of each sample (1-17) are close to each other. Only the isotope values of samples 14a and 14b are deviating from each other. The overall uncertainties depending on the method used range from 0.4 to 0.7‰ (average of overall $\sigma_{f_{15_N}}$ is 0.4‰) for $\delta^{15}N$-NO_3^- and from 1.0 to 1.3‰ for $\delta^{18}O$-NO_3^- (average of overall $\sigma_{g_{18_O}}$ is 1.0‰). The overall uncertainty after correction is composed of the machine error and the error of the sample preparation in the bacterial denitrification method. Since the machine error (0.3‰ for $\delta^{15}N$-N_2O and 0.8‰ for $\delta^{18}O$-N_2O) and the average overall uncertainty (0.4‰ for $\delta^{15}N$-NO_3^- and 1.0‰ for $\delta^{18}O$-NO_3^-) are known, the bacterial denitrification method error for this example batch can be computed as 0.3‰ for $\delta^{15}N$ and 0.6‰ for $\delta^{18}O$.

Table 5-5: Corrected $\delta^{15}N$ (f_1 - f_4) and $\delta^{18}O$ (g_1 - g_2) values and respective uncertainties (σ) of the example batch; the final $\delta^{15}N$ and $\delta^{18}O$ values and overall uncertainties are f_{15_N}, g_{18_O}, $\sigma_{f_{15_N}}$ and $\sigma_{g_{18_O}}$.

Sample	f_1	σ_{f1}	f_2	σ_{f2}	f_3	σ_{f3}	f_4	σ_{f4}	f_{15_N}	$\sigma_{f_{15_N}}$	g_1	σ_{g1}	g_2	σ_{g2}	g_{18_O}	$\sigma_{g_{18_O}}$
Sample 1a	11	0.6	11.3	0.5	11.2	0.4	10.8	0.4	**11**	**0.4**	7.9	1.1	9.4	1.1	**8.7**	**1**
Sample 1b	10.8	0.6	11.1	0.6	11.1	0.4	10.7	0.4	**10.9**	**0.4**	8.2	1.1	9.7	1.1	**8.9**	**1**
Sample 2a	3.9	0.5	4.2	0.5	4.1	0.4	3.7	0.4	**4**	**0.4**	21.2	1.1	23.3	1.2	**22.3**	**1**
Sample 2b	4	0.5	4.3	0.5	4.3	0.4	3.9	0.4	**4.2**	**0.4**	21.6	1.1	23.6	1.2	**22.6**	**1**
Sample 3a	13.5	0.7	13.8	0.6	13.6	0.4	13.2	0.4	**13.5**	**0.4**	8.2	1.1	9.7	1.1	**9**	**1**
Sample 3b	14	0.7	14.2	0.6	14.6	0.4	14	0.4	**14.2**	**0.4**	7.9	1.1	8.1	1.1	**8**	**1**
Sample 4a	11.6	0.6	11.8	0.6	12.3	0.4	11.8	0.4	**11.9**	**0.4**	-9.3	1.2	-9.2	1.1	**-9.2**	**1.1**
Sample 4b	11.8	0.6	11.9	0.6	12.3	0.4	11.8	0.4	**11.9**	**0.4**	-9	1.2	-8.9	1.1	**-9**	**1.1**
Sample 5a	6.7	0.6	6.9	0.6	7.7	0.4	7.1	0.4	**7.1**	**0.4**	12.2	1.1	12.5	1.1	**12.3**	**1**
Sample 5b	6.5	0.6	6.7	0.6	7.7	0.4	7.2	0.4	**7**	**0.4**	14.2	1.1	14.5	1.1	**14.3**	**1**
Sample 6a	13.4	0.7	13.5	0.6	14.1	0.4	13.6	0.4	**13.7**	**0.4**	9.6	1.1	9.9	1.1	**9.7**	**1**
Sample 6b	14	0.7	14.2	0.6	14.5	0.4	13.9	0.4	**14.1**	**0.4**	13.3	1.1	13.6	1.1	**13.5**	**1**
Sample 7a	9.5	0.6	9.7	0.6	10.5	0.4	10	0.4	**10**	**0.4**	7.1	1.1	7.4	1.1	**7.3**	**1**
Sample 7b	9.9	0.6	10	0.6	10.5	0.4	10	0.4	**10.1**	**0.4**	7.5	1.1	7.8	1.1	**7.7**	**1**
Sample 8a	6.4	0.6	6.5	0.6	7.2	0.4	6.7	0.4	**6.7**	**0.4**	13.7	1.1	14	1.1	**13.8**	**1**
Sample 8b	5.9	0.6	6.1	0.5	6.7	0.4	6.1	0.4	**6.2**	**0.4**	13.8	1.1	14.1	1.1	**14**	**1**
Sample 9a	12.9	0.7	13.1	0.6	13.7	0.4	13.2	0.4	**13.2**	**0.4**	8	1.1	8.2	1.1	**8.1**	**1**
Sample 9b	13.2	0.7	13.4	0.6	13.9	0.4	13.3	0.4	**13.5**	**0.4**	8.9	1.1	9.1	1.1	**9**	**1**
Sample 10a	12.5	0.7	12.7	0.6	13.4	0.4	12.9	0.4	**12.9**	**0.4**	7.9	1.1	8.1	1.1	**8**	**1**
Sample 10b	12.3	0.6	12.4	0.6	12.8	0.4	12.2	0.4	**12.4**	**0.4**	9	1.1	9.2	1.1	**9.1**	**1**
Sample 11a	8.4	0.6	9.2	0.6	9.9	0.4	9.2	0.4	**9.2**	**0.4**	5.9	1.1	6.5	1.1	**6.2**	**1.1**
Sample 11b	8.5	0.6	9.2	0.5	9.8	0.4	9.1	0.4	**9.2**	**0.4**	5.3	1.1	5.9	1.1	**5.6**	**1.1**
Sample 12a	5.1	0.6	5.8	0.5	6.5	0.4	5.8	0.4	**5.8**	**0.4**	16.4	1.1	17.1	1.2	**16.7**	**1**
Sample 12b	5.1	0.6	5.9	0.5	6.5	0.4	5.8	0.4	**5.8**	**0.4**	16.9	1.1	17.6	1.2	**17.3**	**1**
Sample 13a	11.7	0.6	12.4	0.6	13.1	0.4	12.4	0.4	**12.4**	**0.4**	6.6	1.1	7.2	1.1	**6.9**	**1.1**
Sample 13b	11.4	0.6	12.1	0.6	12.8	0.4	12.1	0.4	**12.1**	**0.4**	7.1	1.1	7.7	1.1	**7.4**	**1.1**
Sample 14a	5.9	0.6	6.6	0.5	7.5	0.4	6.8	0.4	**6.7**	**0.4**	3.5	1.1	4	1.1	**3.8**	**1.1**
Sample 14b	3.9	0.6	4.7	0.5	5.5	0.4	4.7	0.4	**4.7**	**0.4**	-19.6	1.3	-19.4	1.2	**-19.5**	**1.2**
Sample 15a	8.5	0.6	9.3	0.6	9.9	0.4	9.2	0.4	**9.3**	**0.4**	9.9	1.1	10.6	1.1	**10.3**	**1**
Sample 15b	7.8	0.6	8.5	0.6	9.5	0.4	8.8	0.4	**8.7**	**0.4**	4.9	1.1	5.5	1.1	**5.2**	**1.1**
Sample 16a	8.5	0.6	9.3	0.6	10.2	0.4	9.5	0.4	**9.4**	**0.4**	4.9	1.1	5.5	1.1	**5.2**	**1.1**
Sample 16b	8.6	0.6	9.4	0.6	10	0.4	9.3	0.4	**9.3**	**0.4**	6.5	1.1	7.1	1.1	**6.8**	**1.1**
Sample 17a	3	0.5	3.8	0.5	4.4	0.4	3.7	0.4	**3.7**	**0.4**	19.6	1.1	20.4	1.2	**20**	**1**
Sample 17b	3.2	0.5	3.9	0.5	4.3	0.4	3.6	0.4	**3.8**	**0.4**	19.8	1.1	20.6	1.2	**20.2**	**1**

We analyzed 33 bacterial batches for 561 (33 batches x 17 samples per batch) surface water sample measurements and the computed bacterial denitrification method errors of these 33 batches varied from 0.2 to 2.1‰ for $\delta^{15}N\text{-}NO_3^-$ and from 0.7 to 2.3‰ for $\delta^{18}O\text{-}NO_3^-$, which is much wider than the machine error, which ranges from 0.2 to 0.6‰ for $\delta^{15}N\text{-}N_2O$ and from 0.4 to 1.0‰ for $\delta^{18}O\text{-}N_2O$. The overall uncertainties for the 33 batches range from 0.3 to 2.2‰ for $\delta^{15}N\text{-}NO_3^-$ and from 0.8 to 2.5‰ for $\delta^{18}O\text{-}NO_3^-$. Figure 5-1 demonstrates the overall uncertainties of $\delta^{15}N$ and $\delta^{18}O$ of all the 561 surface water sample measurements (including replicates of each individual sample) in function of the period of analysis.

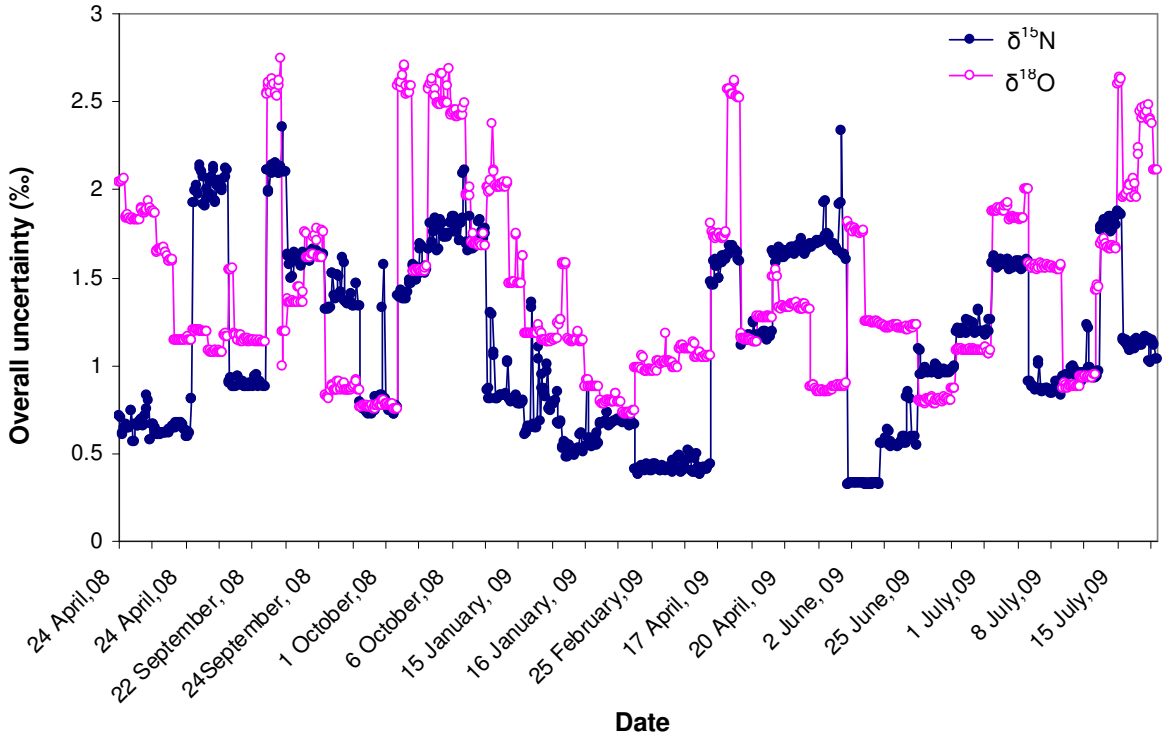


Figure 5-1: The overall uncertainties of $\delta^{15}N$ and $\delta^{18}O$ of all the 561 surface water sample measurements (including replicates of each individual sample) in function of measuring period.

It is obvious that the majority of the overall uncertainty is lower than 1.5‰ for $\delta^{15}N$ and lower than 2‰ for $\delta^{18}O$, excluding some measurements in a certain batch with relatively higher overall uncertainties. The random variability of the overall uncertainty is not related to measuring date, but depending on variations in sample preparation, analytical conditions and equipment functioning for the corresponding batch. Furthermore, in our study, the mean corrected $\delta^{15}N$ and $\delta^{18}O$ values of 132 (4 KNO_3-IWS per batch) KNO_3-IWS measurements in

these batches was 8.4 ± 1.0‰ for $\delta^{15}N\text{-}NO_3^-$ and 25.1 ± 2.0‰ for $\delta^{18}O\text{-}NO_3^-$, which is 1.5‰ lower and 1.1‰ higher for $\delta^{15}N$ and $\delta^{18}O$, respectively, as determined via TC/EA-IRMS. This demonstrates the same tendency of underestimation for $\delta^{15}N$ and overestimation for $\delta^{18}O$ in the bacterial denitrification method as reported in chapter 4. The latter offset for $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ are, however, in the range of the computed overall uncertainties for the 33 batches (0.3 to 2.2‰ for $\delta^{15}N$ and 0.8 to 2.5‰ for $\delta^{18}O$).

5.5 Conclusions

Final corrected $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ values determined by the bacterial denitrification method are derived from a combination of several other measurements of international references and blanks. Each step in the process potentially generates a relatively small uncertainty which propagates through the calculations to yield an overall uncertainty on the final corrected isotope values. The overall uncertainty of the bacterial denitrification method allows using this method for source identification of NO_3^-. However, a slight offset compared to true values is observed.

Chapter 6:

A Bayesian isotope mixing model to estimate proportional contributions of multiple nitrate sources in surface water

This chapter has been edited from:

Xue, D., De Baets, B., Van Cleemput, O., Hennessy, C., Berglund, M. and Boeckx, P. A Bayesian isotope mixing model to estimate proportional contributions of multiple nitrate sources in surface water. Submitted to Environmental Pollution.

6.1 Abstract

In this chapter, we demonstrate a dual isotope approach (δ^{15}N- and δ^{18}O-NO_3^-) and a Bayesian isotope mixing model (SIAR) to identify different NO_3^- sources in surface water and to estimate their proportional contribution. Six sampling points were selected from classes A (agriculture), G (greenhouses in an agricultural area) and H (households). Water samples were collected on a monthly basis from October 2007 to September 2008. An additional aim of our sampling approach was to point out the difference in source contributions between the winter and summer periods. Mean δ^{15}N-NO_3^- values for A, G and H class were given by 15.0-19.4‰, 8.0-12.5‰, and 11.0-15.8‰, respectively, while mean δ^{18}O-NO_3^- values for the A, G and H classes were given by 6.1-13.0‰, 13.7-30.7‰, and 4.5-11.4‰, respectively. SIAR was used to estimate the proportional contribution of five potential NO_3^- sources (NO_3^- in precipitation, NO_3^- fertilizer, NH_4^+ in fertilizer and rain, soil N, and manure and sewage). The analysis showed that "manure and sewage" contributed 32-49% in winter, "soil N", "NO_3^- fertilizer" and "NH_4^+ in fertilizer and rain" contributed 5-28%, while "NO_3^- in precipitation" contributed 2-8%,. In summer, the class A showed similar source contribution patterns as in winter. All five sources contributed more or less equally for the classes G and H in summer (ca. 20%). The advantage of SIAR is that it allows incorporating sources of uncertainty, isotope fractionation and multiple NO_3^- sources. The results confirm that SIAR is an interesting "fingerprint" tool to estimate multiple source contributions, although its usefulness is limited by the wide ranges of isotope values of mixtures and sources.

6.2 Introduction

Nitrate (NO_3^-) contamination in water is an environmental problem worldwide, and is attributed to anthropogenic activities including intensive agriculture, use of fertilizers and animal manure, and discharge of human sewage. To evaluate and manage water quality, NO_3^- concentration monitoring is a widely used approach. However, NO_3^- concentration data alone cannot fully assess the sources and respective contributions of NO_3^- inputs in water, which are key factors in effective management strategies. Although the implementation of the Nitrate Directive (EC, 2002) in Europe established a detailed framework for reducing NO_3^- input to water, NO_3^- is still one of the major contaminants of water resources. Since different NO_3^- sources (fertilizer, manure, human sewage, soil N and atmospheric N deposition) have distinct

isotope ratios of nitrogen ($^{15}N/^{14}N$) and oxygen ($^{18}O/^{16}O$), it is possible to identify these different sources using isotope fingerprints.

A dual isotope approach ($\delta^{15}N$- and $\delta^{18}O$-NO_3^-) can provide meaningful insights for tracing sources of NO_3^- in water. A comparison of 16 watersheds in the U.S. by Mayer et al. (2002), demonstrated that the isotope signature of NO_3^- differed between forested catchment and agricultural land. In predominantly forested watersheds, NO_3^- was mainly derived from soil nitrification processes in soils, resulting in low $\delta^{15}N$-NO_3^- values (less than 5‰). Enriched $\delta^{15}N$ values (between 5 and 8‰) were found in predominantly agricultural watersheds with manure and sewage as contributors. Manure and sewage are enriched in ^{15}N as ammonia (NH_3) volatilization causes an enrichment of ^{15}N in the residual NH_4^+ that is subsequently converted into ^{15}N-enriched NO_3^-. Pardo et al. (2004) successfully used $\delta^{15}N$- and $\delta^{18}O$-NO_3^- to identify atmospheric deposition and microbial nitrification as two main sources of NO_3^- in streams of forested watersheds, as $\delta^{18}O$ signatures of atmospheric NO_3^- (from 25‰ to 75‰) and microbially-produced soil NO_3^- (from 0‰ to 15‰) differ significantly (Xue et al. 2009).

Some researchers also applied $\delta^{15}N$- and $\delta^{18}O$-NO_3^- to quantify different NO_3^- source contributions via a mass-balance mixing model (Phillips and Koch, 2002). Deutsch et al. (2006) successfully used the dual isotope approach to quantify riverine NO_3^- sources, which derived mostly from drainage water (86%), groundwater (11%) and from atmospheric deposition (3%). The dual isotope approach was also applied to quantify NO_3^- contributions into 12 Baltic rivers by Voss et al. (2006). In the study, a mass-balance mixing model was used to quantify three major NO_3^- source contributions, which were sewage, atmospheric deposition and pristine soils. However, a mass-balance mixing model is often performed to find unique solutions with the assumption that there is no variability within sources. In fact, three processes can introduce uncertainty on NO_3^- source apportionment: (a) temporal and spatial variability in $\delta^{15}N$ and $\delta^{18}O$ of NO_3^-; (b) isotope fractionation during denitrification; and (c) too many NO_3^- sources (number of sources > number of isotopes + 1) contribute to the mixture (Moore and Semmens, 2008; Xue et al., 2009).

A Bayesian stable isotope mixing model (Parnell et al., 2010) has been implemented in the software package SIAR (stable isotope analysis in R). This model uses a Bayesian framework to determine the probability distribution of the proportional contribution of each source to a mixture. Furthermore, this mixing model takes into account the uncertainties mentioned above.

The objective of this chapter was to demonstrate that SIAR is a useful "fingerprint" tool for estimating multiple NO_3^- source contributions in complex situations.

6.3 Material and methods

6.3.1 Site description

Flanders is situated in the northern part of Belgium, and about 50% of the total surface area is occupied by agriculture (Cazaux et al., 2007). The vast amount of N present in Flemish surface waters is assumed to originate from intensive manure and mineral fertilizer application in agriculture. Six sampling points from ditches were selected from the MAP network for this study (Figure 6-1): A1 (24050$^{\#}$)(51°12'44" N, 3°35'36" E) located in the Polders of the Bruges basin, A2 (861110$^{\#}$) (51°7'36" N, 2°58'45" E) located in the Yzer basin, G1 (263100$^{\#}$) (51°4'43" N, 4°30'35" E) located in the Nete basin, G2 (376220$^{\#}$) (51°2'59" N, 4°33'12" E) located in the Dijle Zenne basin, H1 (520400$^{\#}$) (50°53'38" N, 3°57'25" E) located in the Dender basin, and H2 (629100$^{\#}$) (50°55'58" N, 3°1'14" E) located in the Leie basin. A1 and A2 are dominated by agricultural inputs; G1 and G2 are dominated by greenhouse inputs in agriculture; and H1 and H2 are dominated by household inputs. During a 10-year monitoring period (1999-2008), annual mean surface water temperature of these six sampling points varied in a similar range from 10 to 16°C and pH values varied between 7 and 8. However, the annual mean NO_3^- concentrations of the six sampling points varied differently: from 5 to 10 mg NL^{-1} for A1, from 10 to 29 mg NL^{-1} for A2, from 24 to 80 mg NL^{-1} for G1, from 15 to 36 mg NL^{-1} for G2, from 2 to 4 mg NL^{-1} for H1 and from 3 to 6 mg NL^{-1} for H2. Fertilizers were yearly applied between February 15 and August 3.

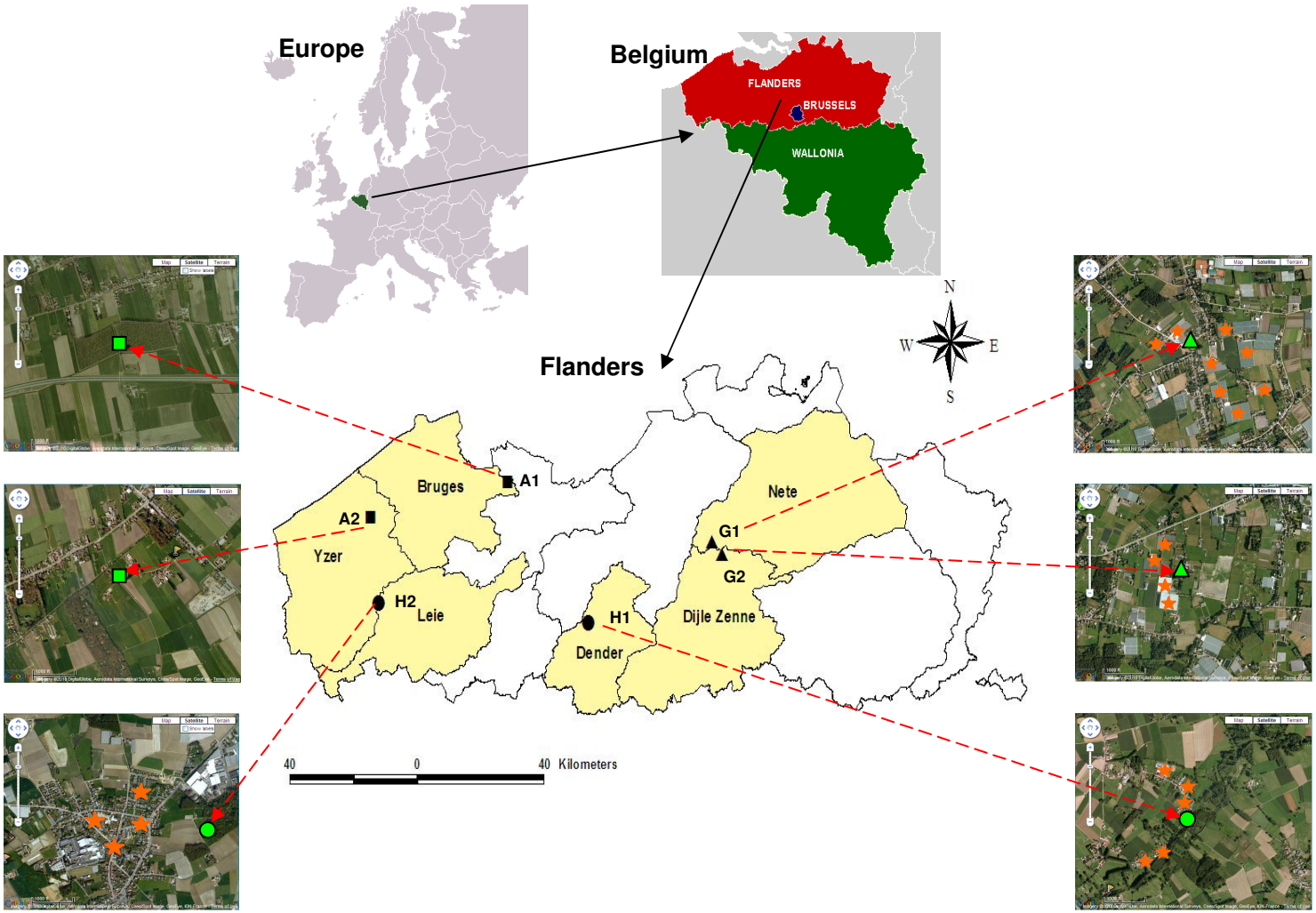


Figure 6-1: Location of the 6 sampling points in the corresponding water basins of Flanders, Belgium. Agriculture (A, ■): 24050# (A1) and 86110# (A2); greenhouses in agriculture (G, ▲): 263100# (G1) and 376220# (G2); and households (H, ●): 520400# (H1) and 629100# (H2). The stars in G1 and G2, and in H1 and H2 represent greenhouses and households, respectively.

6.3.2 Surface water sampling, physico-chemical parameters and isotope analysis

Surface water was sampled monthly from October 2007 to September 2008. Samples were collected in 1L polyethylene bottles and stored in a freezer before analyzing $\delta^{15}N$ and $\delta^{18}O$-NO_3^-. In situ measurement included temperature (T), electrical conductivity (EC 20), pH and dissolved oxygen (DO). Laboratory analyses included NO_3^-, NO_2^-, NH_4^+, Cl^- and PO_4^{3-}. All samples were filtered through 0.45µm membrane filters and stored at 4°C until analysis. Nitrate (NO_3^-) NO_2^-, NH_4^+, and PO_4^{3-} concentrations were analyzed on a Bran + Luebbe Auto Analyzer 3 continuous flow spectrophotometer. Cl^- analyses were carried out by ion chromatography (ion chromatograph 761 Compact IC). The $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values were determined by the "Bacterial denitrification method" (Sigman et al., 2001; Casciotti et al., 2002; Xue et al., 2010), more detailed information can be found in Chapter 5.

In the bacterial denitrification method, the conversion of NO_3^- to N_2O represents a mass balance reaction for N but not for O. Only 1 of the 6 oxygen atoms present in the initial NO_3^- pool is present in the N_2O produced. Furthermore, exchanges of O between nitrogen oxide intermediates and water during the conversion can bias the O isotopic signal of N_2O. Therefore, three international references, USGS32 (180.0 ± 1.0 for $\delta^{15}N$, 25.7 ± 0.4 for $\delta^{18}O$), USGS34 (-1.8 ± 0.2 for $\delta^{15}N$, -27.8 ± 0.4 for $\delta^{18}O$) and USGS35 (2.7 ± 0.2 for $\delta^{15}N$, 56.8 ± 0.3 for $\delta^{18}O$) were used to correct the raw $\delta^{15}N$ and $\delta^{18}O$-NO_3^- values based on a blank correction method and a correction factor method. Details are provided in Chapter 5.

6.3.3 Statistics

To check whether there was a significant difference between $\delta^{15}N$ and $\delta^{18}O$ values among different sample locations, the Tukey HSD (Tukey Honest Significant Difference) test was applied to perform a multiple comparison.

6.3.4 SIAR mixing model

By defining a set of *N* mixture measurements on *J* isotopes with *K* source contributors, the mixing model can be expressed as follows (Parnell et al., 2010):

$$X_{ij} = \sum_{k=1}^{K} p_k (S_{jk} + c_{jk}) + \varepsilon_{ij} \quad (6\text{-}1)$$

$$S_{jk} \sim N\left(\mu_{jk}, \omega_{jk}^2\right)$$

$$c_{jk} \sim N\left(\lambda_{jk}, \tau_{jk}^2\right)$$

$$\varepsilon_{ij} \sim N\left(0, \sigma_j^2\right)$$

where X_{ij} is isotope value j of the mixture i, with i = 1, 2, 3,…, N and j = 1, 2, 3,…, J; S_{jk} is source value k on isotope j (k = 1, 2, 3,…, K) and is assumed to be normally distributed with mean μ_{jk} and standard deviation ω_{jk}; p_k is the proportion of source k, which needs to be estimated by the SIAR model; c_{jk} is the fractionation factor for isotope j on source k and is assumed to be normally distributed with mean λ_{jk} and standard deviation τ_{jk}; and ε_{ij} is the residual error representing the additional unquantified variation between individual mixtures and is assumed to be normally distributed with mean 0 and standard deviation σ_j.

Bayesian techniques allow for the estimation of posterior probability distributions for all f_k through numerical integration. This numerical integration requires randomly generating q proposed vectors of proportional source contributions f_q representing possible states of nature, where all f_k elements in f_q sum to unity. Based on Bayes theorem, the probability of each f_q is then calculated based on data and prior information (Hilborn & Mangel 1997; Ellison 2004):

$$P(f_q | data) = \frac{L(data | f_q) \times p(f_q)}{\sum L(data | f_q) \times p(f_q)} \quad (6\text{-}2)$$

where $L(data | f_q)$ is the likelihood of the data given f_q, $p(f_q)$ is the prior probability of the given state of nature being true based on prior information and $\sum L(data | f_q) \times p(f_q)$ is a numerical approximation of the marginal probability of the data. The numerator $L(data | f_q) \times p(f_q)$ yields the absolute probability of a given f_q based on data and prior beliefs.

A natural prior distribution for f_k is the Dirichlet distribution, often denoted Dir(α_1 ,...,α_k), which treats each source input as independent, but requires summation to unity. The marginal distributions of a Dirichlet distribution (Evans et al., 2000; Jackson et al., 2009) can be explored by defining:

$$\alpha_T = \sum_{k=1}^{K} \alpha_k \tag{6-3}$$

and then the characteristics of the distribution are given by:

$$E[f_k] = \frac{\alpha_k}{\alpha_T} \tag{6-4}$$

$$\mathrm{var}(f_k) = \frac{\alpha_k(\alpha_T - \alpha_k)}{\alpha_T^2(\alpha_T + 1)} \tag{6-5}$$

$$\mathrm{cov}(f_k, f_l) = -\frac{\alpha_k \alpha_l}{\alpha_T^2(\alpha_T + 1)} \tag{6-6}$$

where f_k and f_l (with corresponding Dirichlet parameters α_k and α_l) are the k^{th} and l^{th} source proportions. Since default the SIAR model assumes each α equals 1, each source has prior mean *1/K* and prior variance *(K–1)/(K²(K+1))*. A detailed description of this model can be found in Moore and Semmens (2008), Jackson et al. (2009) and Parnell et al. (2010).

To estimate the contribution of the NO_3^- sources in the 6 sampling points, two isotopes (j = 2) ($\delta^{15}N$- and $\delta^{18}O$-NO_3^-) and five (K = 5) potential NO_3^- sources (NO_3^- in precipitation (NP), NO_3^- fertilizer (NF), NH_4^+ in fertilizer and rain(NFR), soil N (Soil) and manure and sewage (M&S)) were considered in this study. The five potential NO_3^- source values were obtained from literature (reviewed in Xue et al., 2009). The Kolmogorov-Smirnov test for normality was applied to verify whether the literature isotope data of the five NO_3^- sources are normally distributed. Results showed that the $\delta^{15}N$ and $\delta^{18}O$ values of all sources are normally distributed except for $\delta^{15}N$ values of manure and sewage source. Isotope measurements of each sampling point were divided into winter (October to next March) and summer (April to September) measurements for SIAR analysis, since different seasonal source contributions were expected.

Denitrification is a process that results in an exponential increase of $\delta^{15}N$ and $\delta^{18}O$ in NO_3^- as NO_3^- concentration decreases. Some studies reported that this process causes $\delta^{15}N$ and $\delta^{18}O$ to increase in roughly a 2:1 ratio that gives evidence for denitrification (Aravena and Robertson, 1998; Mengis et al., 1999; Fukada et al., 2003). The observed linear relationship between the $\delta^{15}N$- and $\delta^{18}O$ values of the six sampling points implied that no obvious denitrification occurred during the sampling period (data is shown in *6.4.3.1*). The seasonal mean oxygen concentration of all surface waters was above 3.1 mg L^{-1}, which is not ideal for denitrification (Piña-Ochoa and Álvarez-Cobelas, 2006). Hence, corresponding experiments for determining enrichment factors of denitrification were not conducted in this study. Thus, we assumed $\varepsilon_{jk} = 0$ in Eq. (6-1).

6.4 Results and discussion

6.4.1 Physico-chemical data for different land use types

The seasonal physico-chemical data of the six sampling points are summarized in Tables 6-1 and 6-2. The NO_3^- concentrations (Table 6-1) of these sampling points varied widely during the monitoring period, ranging from 0.1 to 78.7 mg N L^{-1}. Among those samples, mean NO_3^--N concentrations for class G were highest (between 19.7 and 41.2 mg N L^{-1}) and showed relatively high standard deviations. By contrast, class H had relatively low mean NO_3^--N concentrations (from 1.3 to 5.0 mg N L^{-1}) and low standard deviations. The mean NO_3^--N concentrations for class A ranged from 1.8 mg N L^{-1} to 18.8 mg N L^{-1}.

The physico-chemical parameters in Table 6-2 show high dissolved O_2 concentrations (3-10 mg O_2 L^{-1}, not favorable for denitrification), high EC20 (electric conductivity which reflects the amount of dissolved ions) and a natural pH (mean pH between 7.1 and 7.8). The seasonal average surface water temperature varied from 6.4 to 8.3°C in winter and from 14.6 to 16.8°C in summer. The sampling points A1, A2, G1, G2 and H1 showed similar ranges for Cl^-, NH_4^+ and NO_2^- . However, H2 showed elevated values, possibly resulting from domestic sewage. Increased PO_4^{3-} concentrations were observed for the sampling points G1 and G2, probably due to phosphate fertilizer use in greenhouses, and for H2, probably due to discharge of domestic sewage.

Table 6-1: Summary of NO_3^- concentrations and isotope statistics for 6 sampling points in winter (W) and summer (S)

Sample	Min. NO_3^-	Mean ± SD NO_3^-	Max. NO_3^-	Min. $\delta^{15}N\text{-}NO_3^-$	Mean ± SD $\delta^{15}N\text{-}NO_3^-$	Max. $\delta^{15}N\text{-}NO_3^-$	Min. $\delta^{18}O\text{-}NO_3^-$	Mean ± SD $\delta^{18}O\text{-}NO_3^-$	Max. $\delta^{18}O\text{-}NO_3^-$
	————mg N/L————			————————————————‰————————————————					
A1 (W)	3.9	7.8 ± 2.1	9.4	15.1	17.8 ± 2.9	21.7	2.1	11.6 ± 7.8	25.2
A1 (S)	0.8	1.8 ± 1.1	3.4	16.2	19.2 ± 2.8	22.7	2.8	9.5 ± 3.8	11.5
A2 (W)	14.0	18.8 ± 3.5	23.0	10.4	15.0 ± 4.7	23.4	3.7	6.1 ± 1.4	7.9
A2 (S)	0.1	6.1 ± 5.6	12.0	14.6	19.4 ± 5.6	27.4	7.4	13.0 ± 5.5	20.6
G1 (W)	11.1	23.4 ± 15.8	54.3	4.7	9.7 ± 3.9	15.6	8.9	13.7 ± 5.1	20.0
G1(S)	16.0	41.2 ± 21.5	77.7	9.6	12.5 ± 3.8	19.0	24.0	28.4 ± 5.6	38.3
G2 (W)	13.3	19.7 ± 6.6	29.8	4.7	8.7 ± 2.4	10.7	12.3	19.0 ± 5.5	25.8
G2 (S)	18.2	37.9 ± 22.5	78.7	4.4	8.0 ± 3.1	13.7	10.2	30.7 ± 13.5	49.5
H1 (W)	2.2	4.5 ± 1.3	5.7	7.5	11.0 ± 2.8	14.6	-0.6	5.6 ± 5.6	13.6
H1 (S)	1.1	5.0 ± 5.5	13.0	13.0	14.2 ± 1.7	15.4	3.0	4.5 ± 2.1	6.0
H2 (W)	0.2	4.2 ± 4.2	11.2	8.7	15.8 ± 6.7	26.4	2.0	7.4 ± 4.0	13.3
H2 (S)	0.1	1.3 ± 1.8	4.5	10.7	15.0 ± 3.8	17.8	4.7	11.4 ± 10.9	24.0

Winter period: October 2007 to March 2008

Summer period: April 2008 to September 2008

A stands for agriculture; G stands for greenhouses in agriculture; and H stands for households

Table 6-2: Summary of physico-chemical data for 6 sampling points in winter (W) and summer (S)

Sample code	Cl^- (mgL^{-1})	NH_4^+ ($mgNL^{-1}$)	NO_2^- ($mgNL^{-1}$)	PO_4^{3-} ($mgPL^{-1}$)	DO (mgL^{-1})	EC 20 ($\mu S/cm^{-1}$)	pH	T (°C)
A1 (W)	41.7 ± 4.4	1.0 ± 0.4	0.2 ± 0.1	0.2 ± 0.1	7.2 ± 2.0	574.8 ± 102.9	7.3 ± 0.2	7.2 ± 2.5
A1 (S)	36.6 ± 4.1	0.7 ± 0.7	0.1 ± 0.1	0.1 ± 0.1	8.9 ± 6.0	589.2 ± 12.3	7.6 ± 0.6	16.8 ± 3.1
A2 (W)	51.3 ± 8.9	0.6 ± 0.4	0.2 ± 0.1	0.2 ± 0.2	9.7 ± 1.8	629 ± 27.1	7.6 ± 0.3	7.7 ± 1.9
A2 (S)	78.8 ± 57.6	5.1 ± 11.2	0.1 ± 0.1	0.6 ± 1.2	7.2 ± 5.4	734.2 ± 293.7	7.4 ± 0.4	15.5 ± 4.1
G1 (W)	46.6 ± 9.4	1.7 ± 1.2	0.2 ± 0.2	1.9 ± 1.9	8.1 ± 1.4	777.8 ± 227.2	7.1 ± 0.2	8.0 ± 3.0
G1(S)	51.6 ± 11.6	1.8 ± 1.7	1.8 ± 1.0	11.6 ± 2.4	7.3 ± 1.2	984.2 ± 274.8	7.2 ± 0.2	16.3 ± 3.9
G2 (W)	37.5 ± 9.7	2.6 ± 2.3	0.2 ± 0.3	7.8 ± 10.7	8.5 ± 2.0	721.3 ± 165.4	7.1 ± 0.2	7.7 ± 2.7
G2 (S)	40.5 ± 7.0	4.0 ± 3.3	0.6 ± 0.4	13.7 ± 6.2	7.4 ± 2.4	885.2 ± 167.8	7.2 ± 0.3	15.1 ± 3.3
H1 (W)	44.4 ± 10.5	1.0 ± 0.5	0.1 ± 0	0.3 ± 0.3	9.9 ± 1.3	663.8 ± 184.8	7.8 ± 0.1	6.4 ± 1.7
H1 (S)	50.8 ± 12.6	2.7 ± 1.9	0.2 ± 0.2	0.5 ± 0.4	4.4 ± 2.7	736.4 ± 114.1	7.5 ± 0.4	14.6 ± 2.9
H2 (W)	165.7 ± 131.5	23.4 ± 18.8	0.4 ± 0.4	4.1 ± 3.9	5.1 ± 2.6	1105.6 ± 678	7.3 ± 0.2	8.3 ± 2.0
H2 (S)	117.5 ± 53.2	41.4 ± 26.5	0.2 ± 0.2	18.2 ± 11.2	3.1 ± 1.5	1150.2 ± 511.9	7.2 ± 0.6	16.0 ± 2.6

Winter period: October 2007 to March 2008

Summer period: April 2008 to September 2008

A stands for agriculture; G stands for greenhouses in agriculture; and H stands for households

6.4.2 Surface water NO_3^- sources for different land use types

The $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of the six sampling points can be found in Table 6-1. The mean $\delta^{15}N$-NO_3^- values were relatively high for class A, and ranged between 15.0‰ and 19.4‰. The $\delta^{15}N$-NO_3^- values in this study are higher than values reported in literature, 11 to 17‰ (Kellman and Hillaire-Marcel, 2003), 10.4‰ (Deutsch et al., 2006), -5 to 11.8 (Lefebvre et al., 2007), and 8.2 to 11.3‰ (Johannsen, 2008). Class G showed relatively low mean $\delta^{15}N$-NO_3^- values, ranging from 8.0‰ to 12.5‰. The mean $\delta^{15}N$-NO_3^- values for class H were intermediate, lying between 11.0‰ and 15.8‰. The mean $\delta^{18}O$-NO_3^- values for class G were highest, from 13.7‰ to 30.7‰. The mean $\delta^{18}O$-NO_3^- values for class H were from 4.5‰ and high to 11.4‰. Class A showed a mean $\delta^{18}O$-NO_3^- range between 6.1 and 13.0‰.

The comparison of the $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of water samples from the three different land use types was summarized in boxplots (Figure 6-2). It is clear that the $\delta^{15}N$-NO_3^- values (Figure 6-2a) of water samples from class A were higher than those of the water samples from classes G and H. Furthermore, a multiple comparison statistical test, Tukey HSD, implied that differences between means of different land use types were all significant at the 95% level ($p < 0.05$). As illustrated in Figure 6-2b, the $\delta^{18}O$-NO_3^- values of water samples from class G were much higher than for the water samples from classes A and H. The Tukey HSD test demonstrated that only the mean of $\delta^{18}O$-NO_3^- values of classes A and H were not significantly different ($p > 0.05$), while the differences between A and G, and G and H were significant at the 95% level ($p < 0.05$).

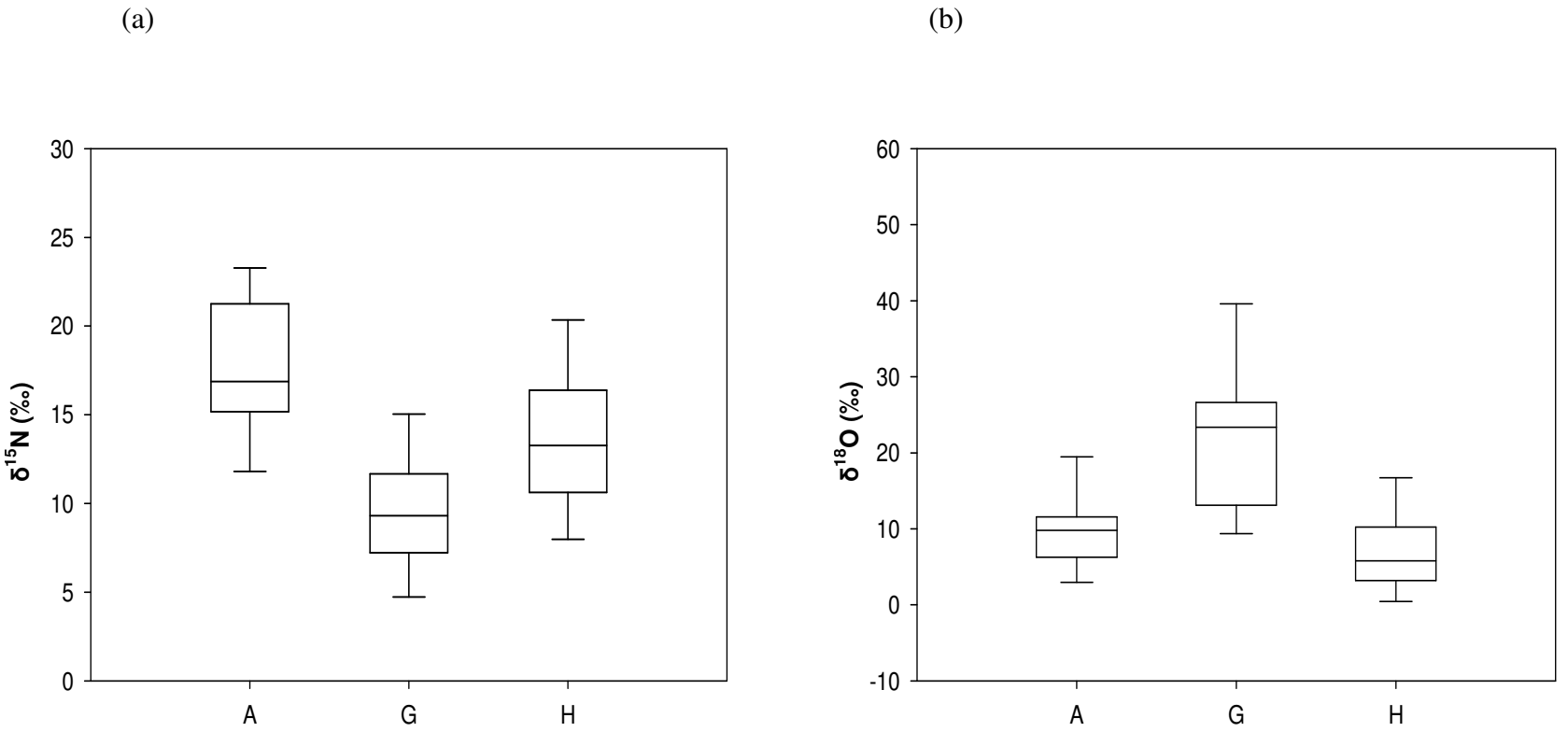


Figure 6-2: Boxplots of δ^{15}N-NO_3^- (a) and δ^{18}O-NO_3^- (b) in expert-knowledge-NO_3^--class during the monitoring period. Boxplots illustrate the 25th, 50th, and 75th percentiles of the amount of water sample isotopic composition; the whiskers indicate the 5th and 95th percentiles of the amount of water sample isotopic composition. A stands for agriculture; G stands for greenhouses in an agricultural area; H stands for households.

A classical dual isotope bi-plot approach (δ^{15}N- NO_3^- vs. δ^{18}O-NO_3^-) was used to identify the predominant NO_3^- source of different sampling points in the three different land use types (Figure 6-3).

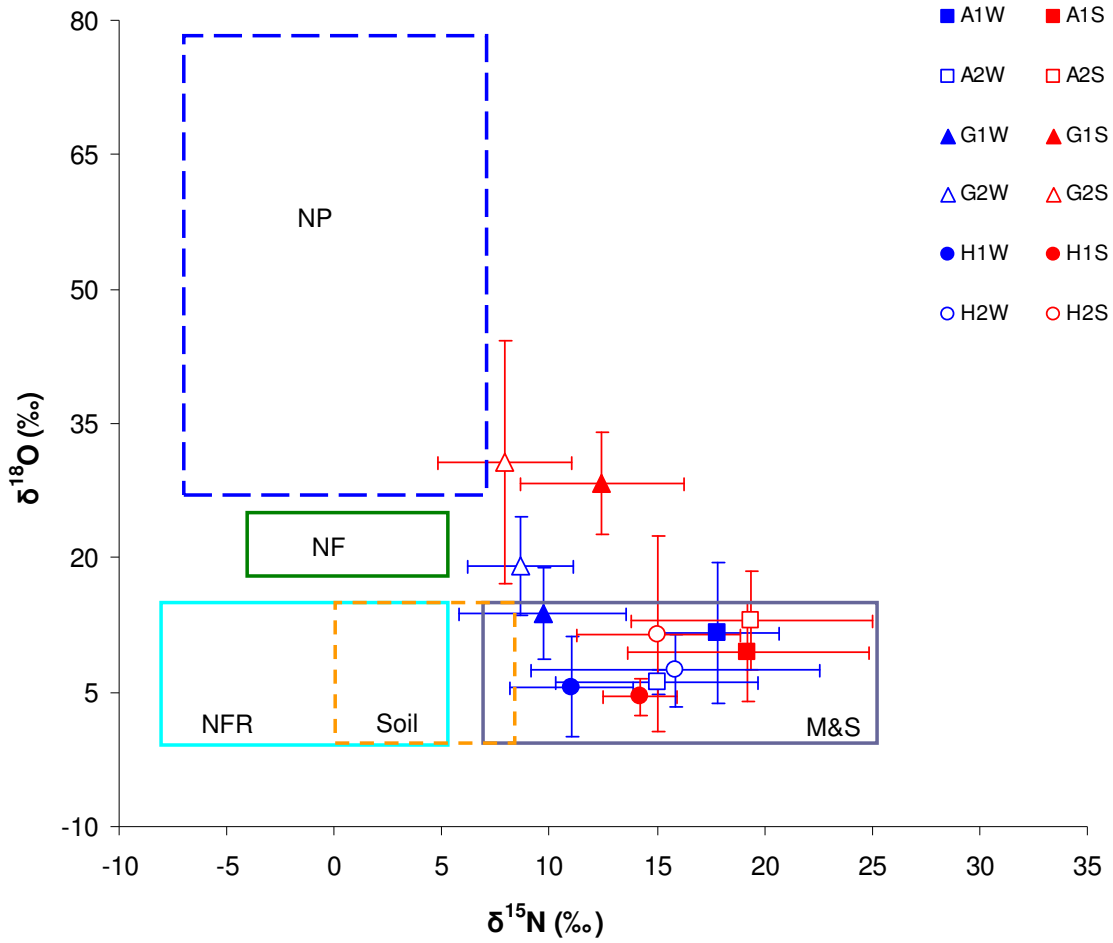


Figure 6-3: Seasonal mean δ^{15}N- and δ^{18}O-NO_3^- values including standard deviation for six sampling points. Ranges of isotopic compositions for five potential NO_3^- sources are determined by Xue et al. (2009) and indicated by boxes: NO_3^- in precipitation (NP), NO_3^- fertilizer (NF), NH_4^+ in fertilizer and rain (NFR), soil N (Soil) and manure and sewage (M&S). A1W-A2W stands for agriculture A1-A2 measured in winter time, while A1S-A2S stands for summer time; G1W-G2W stands for greenhouses in agriculture G1-G2 measured in winter time, while G1S-G2S stands for summer time; H1W-H2W stands for households H1-H2 measured in winter time, while H1S-H2S stands for summer.

It is clear that the isotope signatures of classes A and H mainly fall into the "manure and sewage" source window. For class A, the predominant NO_3^- source could be manure as this is the main fertilizer for crop growth in Flanders (Eppinger et al., 2005). Sewage, however,

could be the predominant NO_3^- source for class H, as the two sampling points were located in the vicinity of sewage discharge points from household areas (Figure 6-1). Furthermore, the contribution from manure cannot be ignored in this area as farmers may feed animals on the grassland nearby the sampling points. Apparent seasonal variation in mean $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of class G could be observed. The mean isotope signatures in summer time are quite close to the "NO_3^- in precipitation" source window. This was likely a result of multiple nitrate source inputs, but with considerable contribution from atmospheric precipitation with high $\delta^{18}O$-NO_3^- in summer season. The mean isotope signatures shifted sharply to the "manure and sewage" source window in winter season. This could result from a significant contribution through agricultural runoff in winter time.

Although we could obtain meaningful information regarding seasonal predominant sources, this is only qualitative information and there is no further quantitative message of the other potential NO_3^- source inputs, which are likely to occur in this complex situation.

6.4.3 Probability estimates of NO_3^- source contributions for different land use types

6.4.3.1 Seasonal variation of different NO_3^- source contributions

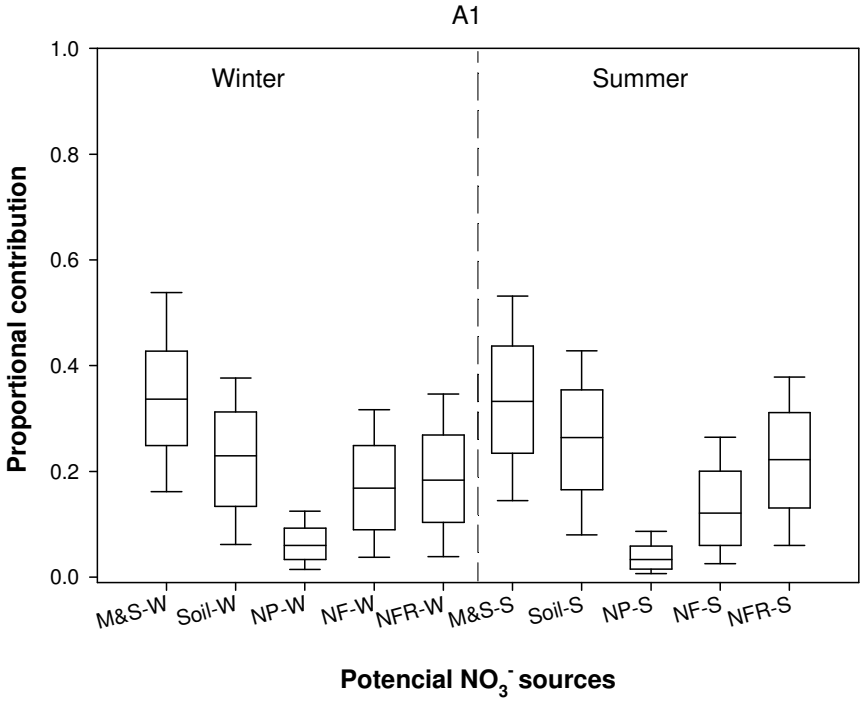
In this study, the observed linear relationship between the $\delta^{15}N$- and $\delta^{18}O$ values of the six sampling points (A1: $\delta^{18}O_{A1} = 0.3\ \delta^{15}N_{A1} + 4.2$; A2: $\delta^{18}O_{A2} = 0.4\ \delta^{15}N_{A2} + 2.7$; G1: $\delta^{18}O_{G1} = 1.0\ \delta^{15}N_{G1} + 8.9$; G2: $\delta^{18}O_{G2} = -0.9\ \delta^{15}N_{G2} + 33.1$; H1: $\delta^{18}O_{H1} = 0.9\ \delta^{15}N_{H1} - 5.1$; H2: $\delta^{18}O_{H2} = -0.6\ \delta^{15}N_{H2} + 18.9$) indicated that no obvious denitrification occurred during the sampling period. Furthermore, measured DO concentrations in the sampling points (mean values > 3.1 mg L^{-1}) are not ideal for denitrification. Therefore, we assumed denitrification to be absent. SIAR was applied to estimate proportional contributions of five potential NO_3^- sources (NO_3^- in precipitation (NP), NO_3^- fertilizer (NF), NH_4^+ in fertilizer and rain (NF&R), soil N and manure and sewage (M&S)) in the three different land use types (A, G and H) for two different seasons (winter and summer) based on $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of water samples from the six sampling points. The $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of local nitrate sources are assumed to be fall in source boxes established from literature data. The isotope values of the five sources are assumed to fall into the source boxes established from literature data (Figure 6-3).

The SIAR mixing model outputs revealed a high variability in contribution of the five potential NO_3^- sources and the ranges of seasonal contributions of each NO_3^- source to the six sampling points are shown in Figure 6-4. Sites A1 and A2 (Figure 6-4a and 6-4b) showed a

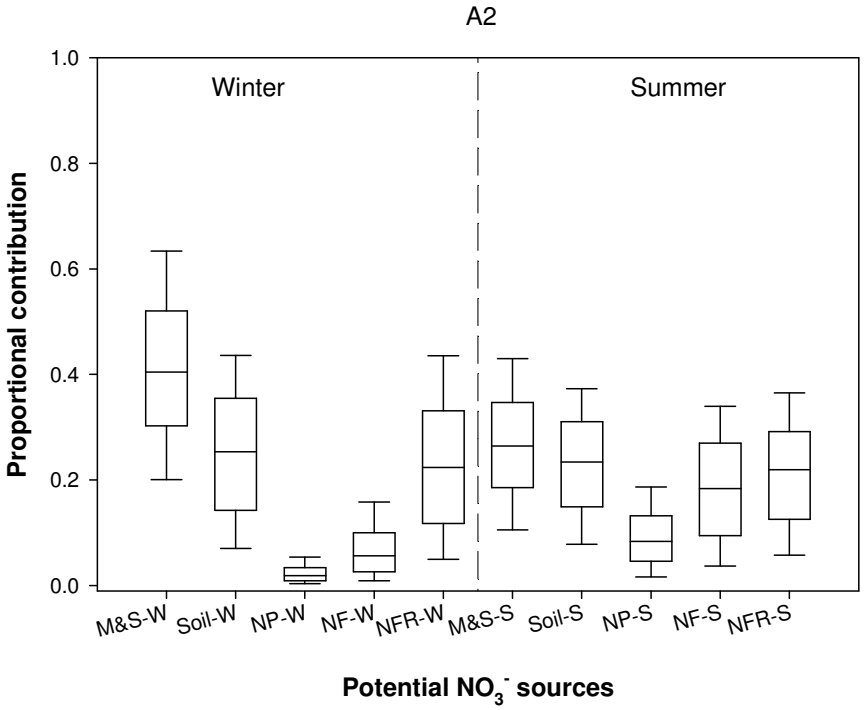
similar source apportionment pattern. The contribution of "M&S" is highest (mean probability estimate between 26% and 40%), followed by "soil N" (mean probability estimate between 23% and 25%), "NF&R" (mean probability estimate between 18% and 23%), "NF" (mean probability estimate 5% and 18%) and "NP" (mean probability estimate between 2% and 8%). Compared to winter time, a decrease of "M&S" source contribution occurred in summer time for A2, which at the same time resulted in an increased contribution of "NF&R". The contribution of "M&S" was almost the same in both seasons for A1. For the two G sites (Figure 6-4c and 6-4d), the source apportionment pattern was slightly different in winter but similar in summer time. For site G1, the source contribution of "M&S" was highest (mean probability estimate equal to 40%), "NP" was lowest (mean probability estimate equal to 8%) and the other three source contributions were intermediate in winter time. For site G2, the "M&S" input was also highest (mean probability estimate equal to 33%) and the other four sources contributed more or less equally (mean probability estimate around 20%). We found that the five potential sources contributed in a similar way to the two G sites in summer. For the G sites that the source of "NP" was higher in summer (mean probability estimate around 30%). Moreover, we also found the highest NO_3^- concentration for the G sites occurred in summer period. One possible explanation could be that farmers with greenhouses collected rainwater in summer to produce fertilizer solutions for crop growth. They then discharged this waste water into surface water after a certain number of re-circulations. The released waste water might lead to high NO_3^- concentrations bearing a mineral fertilizer and rain water signal. The SIAR outputs for the H sites (Figure 6-4e and 6-4f) demonstrated that the two sites had a similar source apportionment pattern. In winter, the contribution of "M&S" was the highest (mean probability estimate from 33% to 49%) and followed by "soil N", "NF&R", "NF" and "NP". In summer, all five sources contribute similarly to the mixture.

In general, SIAR indicated that "M&S" contributed most (mean probability estimate between 32% and 49%) in winter, "NP" contributed least (mean probability estimate between 2% and 8%) in winter except for site G2, and the other three sources, "soil N", "NF" and "NF&R" were intermediate (mean probability estimate between 5% and 28%). In summer, the A sites showed similar source contribution patterns as that in winter time. However, all five sources contributed more or less equally to the G and H sites in summer time (ca. 20%).

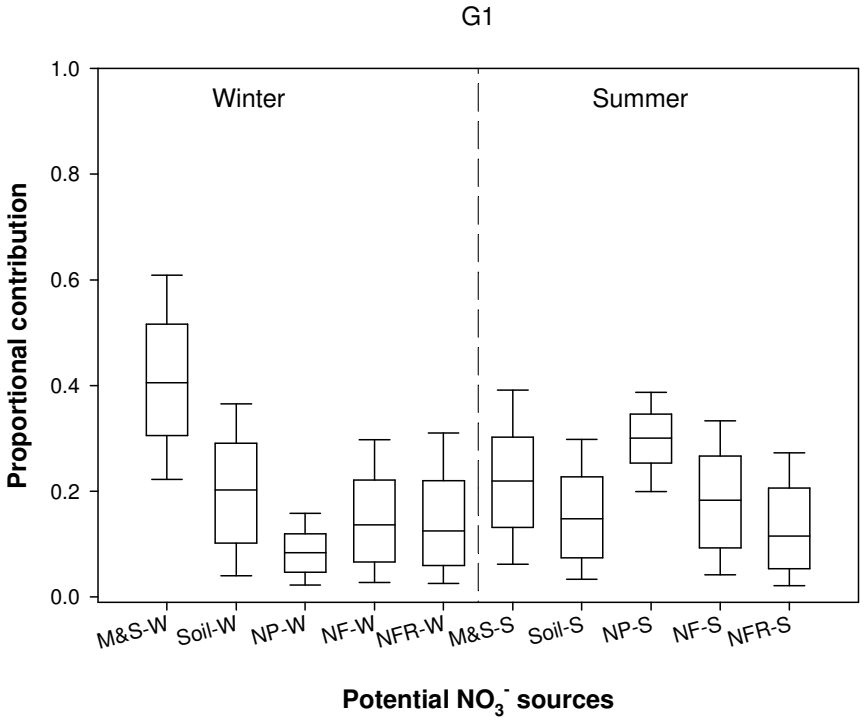
(a)


A1
Winter
Summer
Proportional contribution
1.0
0.8
0.6
0.4
0.2
0.0
M&S-W
Soil-W
NP-W
NF-W
NFR-W
M&S-S
Soil-S
NP-S
NF-S
NFR-S
Potencial NO3- sources

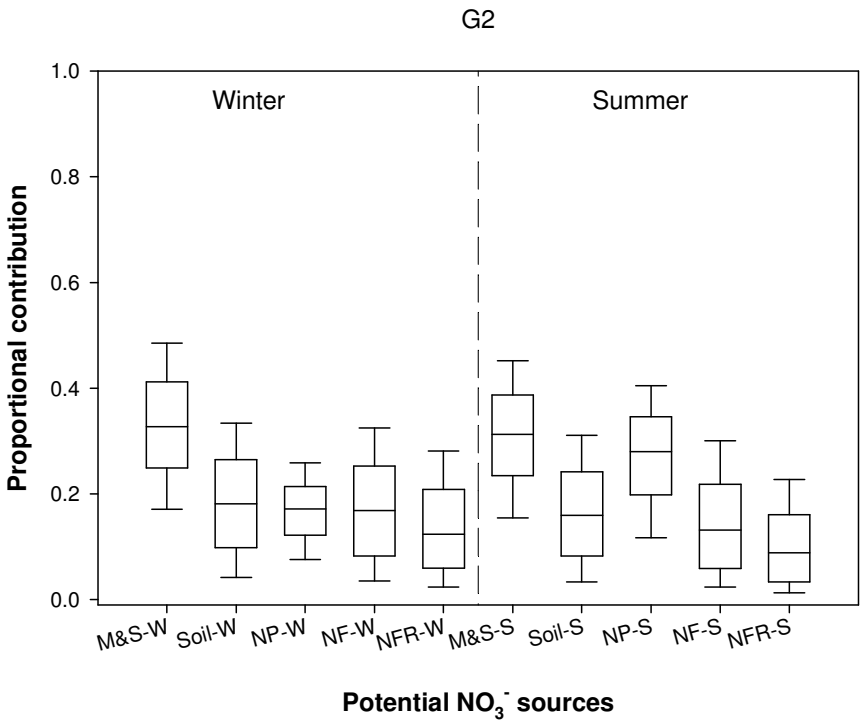
(b)


A2
Winter
Summer
Proportional contribution
1.0
0.8
0.6
0.4
0.2
0.0
M&S-W
Soil-W
NP-W
NF-W
NFR-W
M&S-S
Soil-S
NP-S
NF-S
NFR-S
Potential NO3- sources

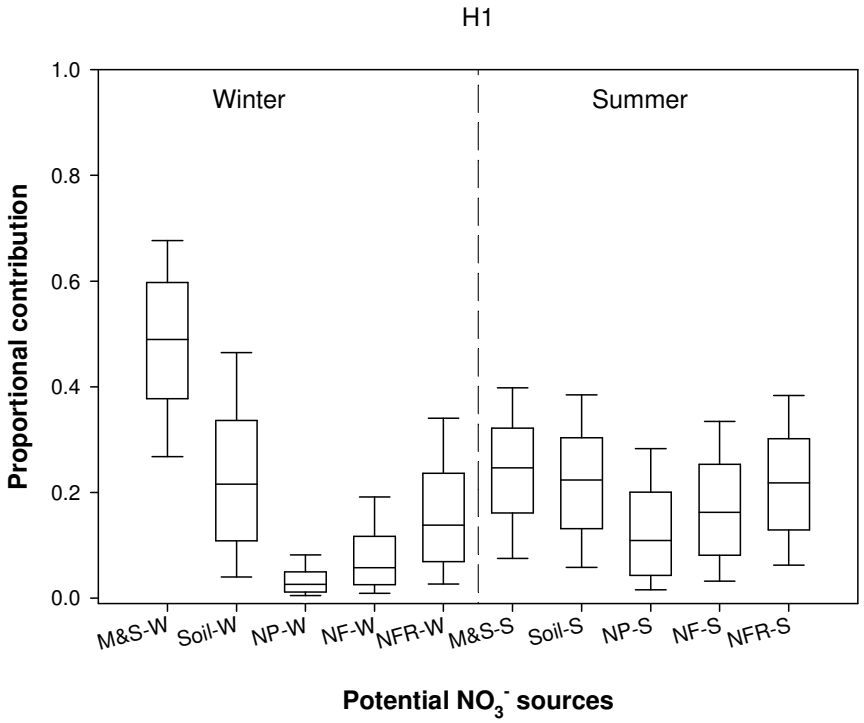
(c)


G1
Winter
Summer
Proportional contribution
1.0
0.8
0.6
0.4
0.2
0.0
M&S-W
Soil-W
NP-W
NF-W
NFR-W
M&S-S
Soil-S
NP-S
NF-S
NFR-S
Potential NO3- sources

(d)


G2
Winter
Summer
Proportional contribution
1.0
0.8
0.6
0.4
0.2
0.0
M&S-W
Soil-W
NP-W
NF-W
NFR-W
M&S-S
Soil-S
NP-S
NF-S
NFR-S
Potential NO3- sources

(e)



(f)

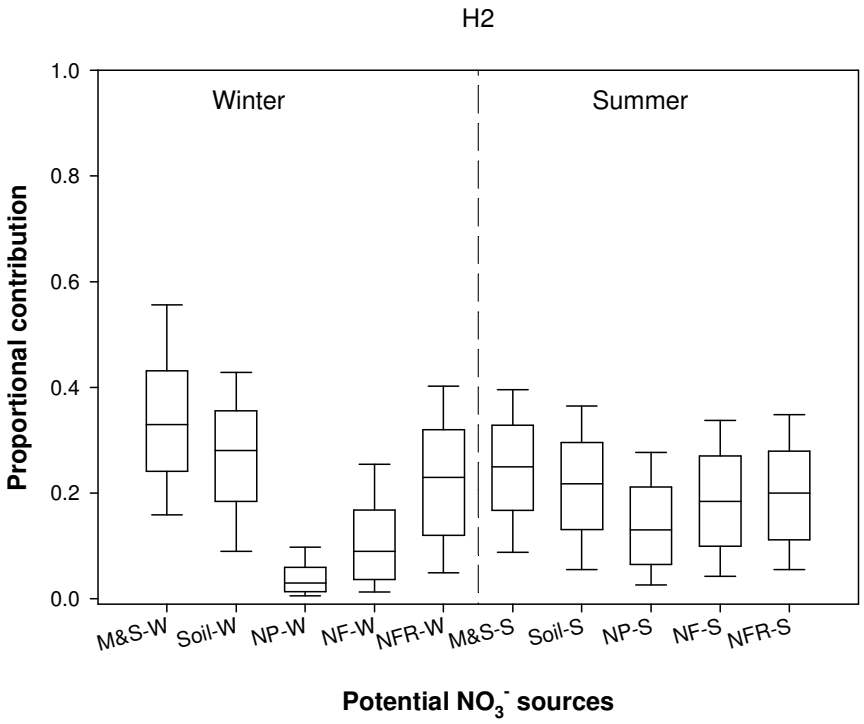


Figure 6-4: Seasonal contributions of five potential NO_3^- sources for six sampling points estimated by SIAR. M&S represents manure and sewage; Soil represents soil N; NP represents NO_3^- in precipitation; NF represents NO_3^- fertilizer and NFR represents NH_4^+ in fertilizer and rain. A stands for agriculture, G stands for greenhouses in agriculture and H stands for households. Boxplots illustrate the 25th, 50th, and 75th percentiles of the estimated source contributions; the whiskers indicate the 5th and 95th percentiles.

In this study, the SIAR output is regarded as a "fingerprint" of potential NO_3^- sources, as it does not only point out the dominant NO_3^- source, but also reveals other important potential NO_3^- sources, which could not be deduced from Figure 6-3. Furthermore, an additional aim of our sampling approach was to lay bare differences in NO_3^- source contributions between summer and winter. The probability estimates of the proportional source contributions (shown in the boxplots in Figure 6-4) therefore include the uncertainty introduced by the temporal variation of the respective NO_3^- sources for a specific sampling location and a given season.

6.4.3.2 Limitations of SIAR for estimation of NO_3^- source contributions

SIAR offers a number of advantages as it can incorporate sources of uncertainty, isotope fractionation and multiple NO_3^- sources. Although this approach is useful, some issues require precaution:

1. $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of the five potential NO_3^- sources have relatively wide ranges and show overlap for NH_4^+ in fertilizer and rain, soil N and manure and sewage (Figure 6-3). Thus, small variations in isotope values of NO_3^- in surface water samples might result in large changes in source apportionment as estimated by SIAR. For example, the seasonal isotope measurements of sampling point H1 are all located in the same manure and sewage source window (Figure 6-3) and the difference of mean isotope values between winter and summer is 3.2‰ for $\delta^{15}N$ and 1.1‰ for $\delta^{18}O$. However, mean probability estimate of this source was 49% in winter and only 24% in summer time, a difference of 25%.
2. The NO_3^- source contribution may vary over time for the six sampling points, which caused the isotopic composition of NO_3^- in water samples to vary as well (see e.g. the large standard deviations in Figure 6-3). Furthermore, the mean $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of the five potential NO_3^- sources also have large standard deviations (Xue et al., 2009). As a result, the SIAR outputs yields a wide range for individual source contributions. The SIAR outputs might be improved by using original source material which could provide more narrow ranges of the isotopic composition of NO_3^- sources. However, this could also constrain NO_3^- source apportionment since source sampling is time consuming and may also show temporal variation in its isotopic compositions.

6.5 Conclusions

Our study showed that a Bayesian mixing model using stable isotope ratios of N and O in NO_3^- could be useful to estimate proportional contributions of NO_3^- sources in surface water. The SIAR output, however, revealed a great variability in contribution of the five potential NO_3^- sources. SIAR provides a "fingerprint" of potential NO_3^- sources, as it does not only demonstrate dominant NO_3^- source contributors, but also reveals other important potential contributors, which cannot be achieved by the classical dual isotope bi-plot approach. In conclusion, SIAR is a useful approach to estimate temporal and spatial variations of different NO_3^- sources. However, its resolution is largely determined by the temporal variability of the isotopic composition of NO_3^- in the mixture and the uncertainty on the isotopic composition of the different NO_3^- sources.

Chapter 7:

Multiple isotope analysis versus expert knowledge for the classification of nitrate sources

This chapter has been edited from:

Xue, D., De Baets, B., Van Cleemput, O., Hennessy, C., Berglund, M. and Boeckx, P. Application of a multiple isotope approach for retrieve of expert knowledge for NO_3^- source classification.

7.1 Abstract

Thirty sampling points with an apriori NO_3^- source classification (6 sampling points per class) based on expert knowledge have been provided: agriculture (class A), agriculture with groundwater compensation (class AGC), a combination of agriculture with horticulture (class AH), greenhouses in an agricultural area (class G), and households (class H). The 30 sampling points were sampled monthly and analyzed for $\delta^{15}N$- and $\delta^{18}O$-NO_3^- during two years (October 2007 to September 2009) and for $\delta^{11}B$ analysis during one year (October 2008 to September 2009). The aim of this study was 1) to retrieve the expert NO_3^- classification via k-means clustering of the outputs of a Bayesian isotope mixing model, 2) compare the performance of two decision tree models using two year physico-chemical data and physico-chemical data with stable isotope data of NO_3^- to correctly classify NO_3^- classes as compared to classification based on expert knowledge and k-means clustering.

A Bayesian stable isotope mixing model (SIAR) has been used to estimate contribution ranges of five potential NO_3^- sources (NO_3^- in precipitation, NO_3^- fertilizer, NH_4^+ in fertilizer and rain, soil N, and manure and sewage) based on $\delta^{15}N$- and $\delta^{18}O$-NO_3^- data. For winter, SIAR estimated "manure and sewage" as major (mean values between 40 and 60%), "NO_3^- in precipitation" as minor (mean values < 10%), and the other three sources, "soil N", "NO_3^- fertilizer" and "NH_4^+ in fertilizer and rain" as intermediate (mean values from 10 to 30%) for classes A, AGC, AH and H. For class G, "manure and sewage" was also the dominant source (mean values ca. 50%) in winter, and the other four sources contributed in a similar range (mean values between 10 and 20%). For summer, as well "manure and sewage" was the dominant source (mean values from 30 to 40%) for classes A and AH as well. The five sources showed a similar pattern in summer as in winter in class AGC. The source contributions of "manure and sewage" and "NO_3^- in precipitation" were dominant (mean values ca. 30%) for class G, and the other three sources contributed in a similar range (mean values between 10 and 20%). For class H, all the five sources contributed in a similar range in summer (mean values ca. 20%). The sampling points in classes A, AGC, AH and G were clustered using k-means clustering. We suggested that 3 clusters were suitable for winter and summer, as it gave the same silhouette value (0.6) and Rand index (0.7) as for 4 clusters. Moreover, the majority of the sampling points stayed in the same class as the expert classification. Since $\delta^{11}B$ showed sewage was the dominant source for class H, this class was not conducted by k-means. In conclusion the sampling points were divided into four classes both in winter and summer: classes A, class AGC, class G, and class H compared to the expert

classification. Thus, the expert class AH seems not to be retained based on isotopic characterization. Comparison of decision tree models built on physico-chemical data and physico-chemical data and stable isotope data of nitrate with expert classification (5 classes) and k-means clustering classification (4 classes) indicated that 1) isotopes help to develop a NO_3^- classification in absence of expert knowledge and 2) decision tree models based on physico-chemical data could be helpful to classify new sampling points in one of the NO_3^- classes.

7.2 Introduction

Nitrate (NO_3^-) pollution from agricultural areas is one of the major anthropogenic sources of NO_3^- in surface waters in Europe (European Environmental Agency, 2005). The implementation of the Nitrate Directive (EC, 2002) in Europe established a detailed framework for the protection of waters due to NO_3^- pollution from agricultural sources and imposed a maximum allowable NO_3^- concentration of 50 mg NO_3^- L^{-1}. Therefore, European countries design monitoring programmes to identify water status, evaluate pressure on water systems, and detect water quality trends. Although monitoring the NO_3^- concentration in water can provide useful information for NO_3^- pollution and its management, this approach requires a dense and long-term monitoring scheme with associated high cost. Furthermore, NO_3^- concentrations alone cannot precisely identify NO_3^- sources responsible for water contamination.

Since different NO_3^- sources (fertilizer, manure, industrial or septic waste, atmospheric N deposition) have unique isotope ratios for nitrogen ($^{15}N/^{14}N$) and oxygen ($^{18}O/^{16}O$), it is possible to identify NO_3^- pollution sources using isotope fingerprints, i.e. discriminating inorganic fertilizer from organic fertilizer (Kendall and Aravena, 1999), microbial nitrification from sewage or manure (Johannsen et al., 2008), forested catchment from agricultural land (Mayer et al., 2002), atmospheric deposition from microbial nitrification (Pardo et al., 2004, Barnes et al., 2008). However, multiple NO_3^- sources from agricultural and urban activities contributing to surface water and groundwater and isotope fractionation caused by multiple N-transformation processes could change the original $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values, potentially biasing identification of NO_3^- contaminated sources (Kendall, 1998; Kellman, 2005; Xue et al., 2009). It has been demonstrated that boron (B) is an additional isotope tracer for identification of NO_3^- sources in water (Bassett et al., 1995; Vengosh et al., 1999; Widory et al., 2004; Widory et al., 2005). Boron has two stable isotopes (^{11}B and ^{10}B) with natural

abundances of approximately 80% and 20%. Boron can co-migrate with NO_3^- and is not affected by transformation processes (Vengosh et al., 1994; Bassett et al., 1995; Leenhouts et al.1998). Widory et al. (2004, 2005) demonstrated the benefits of the combined use of $\delta^{15}N$ and $\delta^{11}B$ values to identify multiple NO_3^- sources (fertilizers, greenhouse discharges, sewage, hog, cattle and poultry manure) in groundwater in areas with different hydrogeological conditions in France. Seiler (2005) used $\delta^{15}N$ and $\delta^{18}O$ values combined with $\delta^{11}B$ values to discriminate between domestic wastewater and fertilizer contamination in groundwater in Nevada (USA).

Quantification of different potential NO_3^- sources would provide further meaningful information for decision makers. Deutsch et al. (2006) and Voss et al. (2006) both applied a mass-balance mixing model (Phillips and Koch, 2002) to quantify different NO_3^- source contributions into water. Recently, a Bayesian stable isotope mixing model (Parnell et al., 2010) has been implemented in the software package SIAR (stable isotope analysis in R). This model uses a Bayesian framework to determine robust estimates of source proportions of each source to a mixture. Furthermore, this mixing model takes into account the incorporation of source uncertainties, which are not considered in a classical mass-balance mixing model, for NO_3^- source apportionment: (a) temporal and spatial variability in $\delta^{15}N$ and $\delta^{18}O$ of NO_3^-; (b) isotope fractionation during denitrification; and (c) too many NO_3^- sources (number of sources > number of isotopes + 1) contribute to the mixture (Moore and Semmens, 2008; Xue et al., 2009; Parnell et al., 2010).

In chapter 3, thirty sampling points were selected from the MAP network for isotope monitoring and were classified into 5 different classes by expert knowledge: agriculture (class A), agriculture with groundwater compensation (class AGC), a combination of agriculture with hotticulture (class AH), greenhouses in an agricultural area (class G), and households (class H).

The aim of this chapter is to (1) apply a multiple isotope approach, $\delta^{15}N$-NO_3^-, $\delta^{18}O$-NO_3^- and $\delta^{11}B$, to identify predominant NO_3^- sources for Flemish surface waters affected by different activities, (2) apply a Bayesian stable isotope mixing model to estimate proportional contributions of potential NO_3^- sources and (3) verify expert classifications of the sampling points using outputs of a Bayesian isotope mixing model in a k-means clustering approach.

7.3 Material and methods

7.3.1 Site description

This study consists of thirty sampling points that were selected from the MAP monitoring network. These sampling points are distributed in different basins over the whole of Flanders (Figure 7-1).

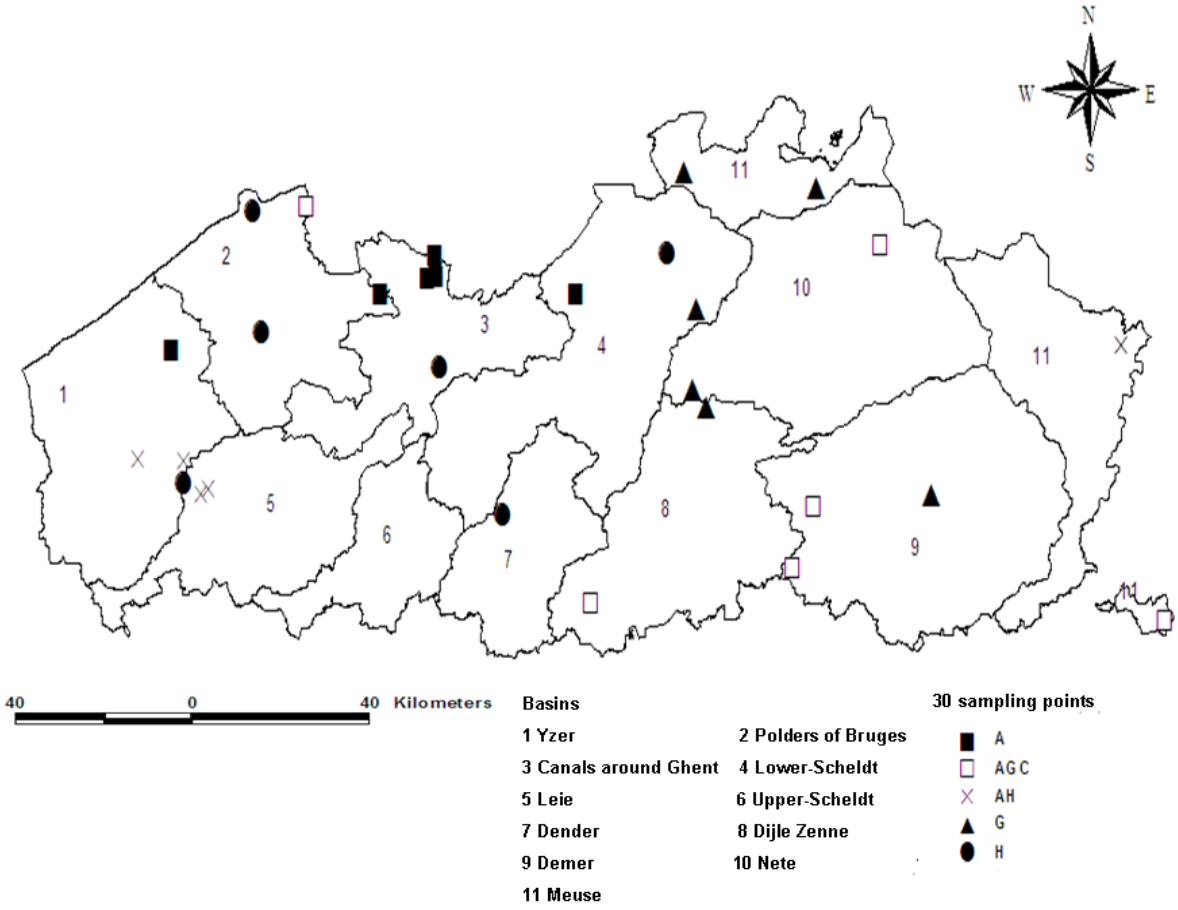


Figure 7-1: Location and classification of 30 sampling points in Flanders. A represents agriculture; AGH represents agriculture with groundwater compensation; AH represents a combination of agriculture with horticulture; G represents greenhouses in an agricultural area; H represents households.

Experts from VMM classified the sampling points into 5 different classes based on different land use types: 6 sampling points in class A (A1-A6), 6 sampling points in class AGC (AGC1-AGC6), 6 sampling points in class AH (AH1-AH6), 6 sampling points in class G (G1-G6), and 6 sampling points in class H (H1-H6). The coordinates of locations and abbreviation of the sampling points numbers are shown in Table 7-1

Table 7-1: Summary information regarding 30 sampling points.

NO_3^- class	Abbreviation	Sampling code	Basin	Coordinate
A	A1	24050	Polders of Bruges	51°12'44" N, 3°35'36" E
	A2	861110	Yzer	51°7'36" N, 2°58'45" E
	A3	13000	Canals around Ghent	51°16'12" N, 3°45'10" E
	A4	13500	Canals around Ghent	51°14'21" N, 3°45'20" E
	A5	17000	Canals around Ghent	51°14'8" N, 3°43'45" E
	A6	194660	Lower-Scheldt	51°12'47" N, 4°10'2" E
AGC	AGC1	26720	Polders of Bruges	51°20'28" N, 3°22'16" E
	AGC2	153400	Meuse	50°43'51" N, 5°52'51" E
	AGC3	307100	Nete	51°17'10" N, 5°3'49" E
	AGC4	365520	Dijle Zenne	50°45'57" N, 4°12'36" E
	AGC5	408760	Demer	50°54'17" N, 4°51'48" E
	AGC6	426520	Demer	50°49'N, 4°48'4" E
AH	AH1	130300	Meuse	51°7'59" N, 5°46'12" E
	AH2	130350	Meuse	51°8'4" N, 5°46'11" E
	AH3	624515	Leie	50°55'31" N, 3°5'36" E
	AH4	624550	Leie	50°55'6" N, 3°4'16" E
	AH5	926100	Yzer	50°57'54" N, 3°1'12" E
	AH6	938210	Yzer	50°57'59" N, 2°53'8" E
G	G1	263100	Nete	51°4'43" N, 4°30'35" E
	G2	376220	Dijle Zenne	51°2'59" N, 4°33'12" E
	G3	65100	Meuse	51°23'40" N, 4°29'11" E
	G4	83900	Meuse	51°22'17" N, 4°52'41" E
	G5	190220	Lower-Scheldt	51°11'41" N, 4°31'27" E
	G6	449890	Demer	50°55'12" N, 5°12'36" E
H	H1	520400	Dender	50°53'38" N, 3°57'25" E
	H2	629100	Leie	50°55'58" N, 3°1'14" E
	H3	6009	Polders of Bruges	51°19'51" N, 3°12'41" E
	H4	57500	Canals around Ghent	51°6'28" N, 3°46'6" E
	H5	183700	Lower-Scheldt	51°16'31" N, 4°26'17" E
	H6	883560	Polders of Bruges	51°9'18" N, 3°14'43" E

7.3.2 Surface water sampling and isotope analysis

Surface water was sampled by VMM monthly from October 2007 to September 2009. Samples were collected in 1L polyethylene bottles and stored in a freezer before analyzing δ^{15}N- and δ^{18}O-NO_3^-. In situ analyses included temperature (T), electrical conductivity (EC 20), pH and dissolved oxygen (DO). Laboratory analyses included NO_3^-, NO_2^-, NH_4^+, Cl^- and PO_4^{3-}. All samples were filtered through 0.45μm membrane filters and stored at 4°C until analysis. The methods for measurement of physico-chemical properties and determination and δ^{15}N- and δ^{18}O-NO_3^- analysis were described in Chapter 6.

During the isotope monitoring period, some sampling points showed relatively low sampling frequency, i.e. A4, A6, AGC1 and AH2, which could result from either dry seasons or

missing samples during transports. Some sampling points displayed different land use types as defined by the experts on googlemap, i.e. G3, G4, and H3. Thus, sampling points with relatively high sampling frequency and in the same land use type were selected from each class for $\delta^{11}B$ analysis from October 2008 to September 2009. Due to the high cost of $\delta^{11}B$ analysis, surface water samples of the selected homogeneous sampling points from one class were mixed together in equal volumes (200 mL) to create one composite sample for $\delta^{11}B$ analysis. Thus, five mixed samples were created, which are class A (A1-A3, A5), class AGC (AGC2-AGC6), class AH (AH1, AH3- AH6), class G (G1-G2, G5-G6), and class H (H1-H2, H4-H6).

The $\delta^{11}B$ values are determined by isotopic ratio measurements following purification and concentration by filtration, ion exchange and sublimation (Lemarchand et al., 2002). Purified samples are loaded onto Ta filaments with cesium carbonate, mannitol and graphite and introduced into a single focusing sector field TIMS (thermal ionization mass spectrometry) instrument (National Bureau of Standards, NBS). The NBS is equipped with a single Faraday collector and a 10^{11} Ω resistor. Ion beams of $(Cs_2{}^{10}BO_2)+$ and $(Cs_2{}^{11}BO_2)+$, with *m/e* equal to 308 and 309 respectively, are measured. All measured isotope ratios are corrected for ^{17}O contribution at *m/e* 309 ($Cs_2{}^{10}B^{16}O^{17}O$). As with N and O, B isotopic ratios are expressed in delta (δ) units and a per mil (‰) notation relative to an international standard, in this case NIST SRM 951 (boric acid).

7.3.3 SIAR mixing model

More detailed information on the SIAR mixing model can be found in Chapter 6. Denitrification is a process that results in an exponential increase of $\delta^{15}N$ and $\delta^{18}O$ in NO_3^- as NO_3^- concentration decreases. Some studies reported that this process causes $\delta^{15}N$ and $\delta^{18}O$ to increase roughly in a 2:1 ratio giving evidence for denitrification (Aravena and and Robertson, 1998; Mengis et al., 1999; Fukada et al., 2003). The observed linear relationship between the $\delta^{15}N$- and $\delta^{18}O$ values of the 30 sampling points implied that no obvious denitrification occurred during the sampling period (data not shown). The overall mean oxygen concentration of these surface waters was 7.0 mg L^{-1}, which is not ideal for denitrification (Piña-Ochoa and Álvarez-Cobelas, 2006). Hence, corresponding experiments for determining enrichment factors of denitrification were not conducted in this study.

7.3.4 K-means clustering

K-means clustering (MacQueen, 1967) is a popular unsupervised learning algorithm. It partitions a data set into *k* clusters. The basic principle is that it finds a partition in which data points within each cluster are as close to each other as possible, and as far as possible from data points in other clusters. Assume there are *N* data points $x_1, x_2, \ldots, x_n$ to be clustered into *k* different subsets C_i, $i=1,\ldots, k$, each containing n_i data points, $0<n_i<N$. K-means clustering minimizes the following mean-squared-error (MSE) cost function (Žalik, 2008):

$$J_{MSE} = \sum_{i=1}^{k} \sum_{x_t \in C_i} \left\| x_t - c_i \right\|^2 \tag{7-1}$$

x_t is a vector representing the *t*-th data point in cluster C_i and c_i is the geometric centroid of cluster C_i. The aim of this function is to minimize the sum of distances between data points x_t and cluster centers c_i. This algorithm is composed of the following steps:

- Step 1: Initialize *k* cluster center positions $c_1, c_2,\ldots, c_k$ by some initial values using random sampling;
- Step 2: Assign each data point to a cluster that contains the closest centroid;
- Step 3: When all data points have been assigned, recalculate the positions of the *k* centroids;
- Step 4: Repeat steps 2 and 3 until the centroids no longer move.

The technique used to measure how similar a data point is to data points in its own cluster compared to data points in other clusters, is called "silhouette validation". The average silhouette value can be used for evaluation the clustering validity and to decide how good the number of selected clusters is. The silhouette value *S(i)* is given by (Rousseeuw, 1987):

$$S(i) = \frac{(b(i) - a(i))}{\max\{a(i), b(i)\}} \tag{7-2}$$

where *a(i)* is the average distance from the *i*-th data point to the other data points in its cluster, and *b(i)* is the minimum average distance from the *i*-th data point to data points in another cluster; where the average is computed for each cluster separately. It follows that $-1 \leq S(i) \leq 1$. A silhouette value close to 1 means that data points are very distant from neighboring clusters. A silhouette value of zero means that data points could probably be assigned to another cluster, and the data points lie equally far away from both clusters. A silhouette value close to

–1 means that points have probably been assigned to a wrong cluster. The overall average silhouette value is simply the average of the *S(i)* for all points in the data set. The largest overall average silhouette value indicates the best clustering. Thus, the corresponding number of clusters is taken as the optimal number of clusters.

7.3.5 Rand Index

The Rand index (Rand, 1971) is a measure for the similarity between two data clustering. Given a set of *n* objects $S = \{O_1, \dots, O_n\}$ and two partitions *P* and *Q* of *S*, with *r* and *c* subsets to compare: $P = \{p_1, \dots, p_r\}$ and $Q = \{q_1, \dots, q_c\}$. We define:

- *a*, the number of pairs of objects that are in the same set in *P* and in the same set in *Q* ;
- *b*, the number of pairs of objects that are in the same set in *P* and in different sets in *Q*;
- *c*, the number of pairs of objects that are in different sets in *P* and in the same set in *Q*;
- *d*, the number of pairs of objects that are in different sets in *P* and in different sets in *Q*;

The Rand index is defined as:

$$\text{Rand} = \frac{a+d}{a+b+c+d} \tag{7-3}$$

where $a + d$ represents the number of agreements between *P* and *Q* and $b + c$ represents the number of disagreements between *P* and *Q*. Thus, the Rand Index is a similarity measure which assumes values between 0 and 1: with (a) values greater than 0.90, an excellent agreement can be considered; with (b) values greater than 0.80, a good agreement can be considered; with (c) values greater than 0.65, a moderate agreement can be considered; with (d) values less than 0.65, a poor agreement is considered.

7.3.6 Decision tree model

The WEKA 3.4.10 software has been used to explore a decision tree. This software uses J4.8 algorithm, which is WEKA's implementation of the C4.5 decision tree learner (Quinlan, 1993). More detailed information for a decision tree model description can be found in Chapter 3.

7.4 Results and discussion

7.4.1 Physico-chemical data for different NO_3^- source classes

The physico-chemical parameters of the 30 sampling points are summarized in Tables 7-2, 7-3 and 7-4. The physico-chemical parameters in Table 7-2 demonstrated that the 30 sampling points were high in dissolved O_2 (relatively high DO concentration which indicates that it is not ideal for denitrification), had a natural pH (mean pH between 6.3 and 7.8) and moderate temperature (T) (mean T from 9.7 to 14.1°C). Most sampling points showed similar ranges of Cl^-, NH_4^+, NO_2^- and PO_4^{3-}. However, exceptionally high values were observed in the parameter of Cl^- for sampling points A3, A4, AGC1, H2 and H3, possibly resulting from overuse of (mineral) fertilizers (i.e. potassium chloride, magnesium chloride, and ammonium chloride fertilizer), or discharge of domestic sewage. In addition, sewage discharge might also cause relatively high NH_4^+ values for sampling points H2 and H4.

Table 7-2: Summary of physico-chemical data for 30 sampling points

Sampling code	Cl^- (mg L^{-1})	NH_4^+ (mg N L^{-1})	NO_2^- (mg N L^{-1})	PO_4^{3-} (mg P L^{-1})	DO (mg O_2 L^{-1})	EC 20 (µS cm^{-1})	pH	T (°C)
A1	37.8 ± 6.4	0.8 ± 0.8	0.1 ± 0.1	0.1 ± 0.1	7.4 ± 3.4	536.6 ± 127.4	7.3 ± 0.4	11.5 ± 6.4
A2	79 ± 49.9	5.2 ± 9.7	0.2 ± 0.2	0.9 ± 1.7	8.0 ± 3.8	745.9 ± 262.1	7.5 ± 0.4	13.0 ± 6.1
A3	8824 ± 309.9	0.8 ± 1.3	0.1 ± 0	0.7 ± 0.8	8.1 ± 2.5	3190 ± 970	7.8 ± 0.4	11.8 ± 5.5
A4	554.4 ± 183.1	2.8 ± 4	0.1 ± 0	1.9 ± 2.1	6.4 ± 3.7	2435.4 ± 534.8	7.7 ± 0.5	11.5 ± 4.7
A5	81.5 ± 22.9	1.1 ± 1	0.1 ± 0.1	0.3 ± 0.3	6.2 ± 3.2	727.7 ± 111.4	7.4 ± 0.3	11.7 ± 5.9
A6	36.9 ± 11.2	0.4 ± 0.5	0.1 ± 0	0.3 ± 0.7	5.8 ± 3.0	776.3 ± 160.3	7.3 ± 0.3	10.6 ± 5.9
AGC1	9427.2 ± 4194.6	0.7 ± 0.3	0.1 ± 0	2.6 ± 3.4	5.5 ± 3.0	17852.5 ± 7508.7	7.6 ± 0.5	11.8 ± 4.9
AGC2	16.1 ± 3.5	0.1 ± 0.1	0.1 ± 0	0.2 ± 0.3	9.5 ± 0.9	615.4 ± 83	7.5 ± 0.4	10.7 ± 2.9
AGC3	36.2 ± 3.1	0.3 ± 0.3	0.1 ± 0	0.2 ± 0.7	8.4 ± 1.6	287.5 ± 20.7	6.7 ± 0.4	11.9 ± 4.9
AGC4	42.5 ± 7.3	1.2 ± 1.5	0.1 ± 0.1	1.7 ± 3.2	7.3 ± 2.5	779.2 ± 70.8	7.6 ± 0.4	11.0 ± 5
AGC5	38.5 ± 8.6	0.3 ± 0.2	0.1 ± 0.1	1.2 ± 2.8	10.5 ± 1.7	324.6 ± 41.4	7.4 ± 0.3	9.9 ± 4.5
AGC6	43.7 ± 3.9	0.2 ± 0.2	0.1 ± 0	0.2 ± 0.3	9.2 ± 1.6	653.4 ± 61.6	7.4 ± 0.4	10.4 ± 2.6
AH1	36 ± 7.2	0.3 ± 0.2	0.1 ± 0	0.3 ± 0.3	6.6 ± 3.4	477.4 ± 145	6.3 ± 0.5	10.1 ± 5
AH3	47.3 ± 8.6	1.1 ± 1.2	0.2 ± 0.1	0.4 ± 0.5	8.5 ± 2.2	636.5 ± 117	7.4 ± 0.5	12.3 ± 4.6
AH4	54.6 ± 30.1	7.9 ± 22.4	0.2 ± 0.2	0.9 ± 2.3	7.8 ± 3.2	715.2 ± 355.4	7.3 ± 0.3	12.6 ± 4.6
AH5	65.8 ± 24.3	3.6 ± 5.4	0.3 ± 0.2	0.4 ± 0.5	7.5 ± 1.6	806.1 ± 134.4	7.3 ± 0.4	11.0 ± 4.4
AH6	46.9 ± 6.9	0.2 ± 0.2	0.1 ± 0.1	0.2 ± 0.2	10.2 ± 2.1	620.1 ± 65.9	7.5 ± 0.5	10.6 ± 4.3
G1	47.7 ± 10.8	2.2 ± 1.5	1 ± 1.5	4.2 ± 4.5	7.8 ± 1.7	857.7 ± 277.9	7.2 ± 0.3	11.2 ± 5.9
G2	40.2 ± 8.8	3.8 ± 2.6	0.4 ± 0.3	6 ± 7.7	8 ± 2.8	741.4 ± 169.2	7.2 ± 0.4	11.3 ± 5
G3	29.7 ± 11.2	0.9 ± 0.5	0.1 ± 0.1	0.3 ± 0.4	5.9 ± 1.5	379.1 ± 56.8	6.4 ± 0.3	10.7 ± 4.9
G4	52.5 ± 45.9	4.8 ± 3.7	0.4 ± 0.5	3.6 ± 5.2	4.8 ± 2.7	542.4 ± 378.6	6.9 ± 0.9	13.8 ± 6.3
G5	31.5 ± 5.3	0.7 ± 0.6	0.1 ± 0.1	2.8 ± 6.6	7.9 ± 1.9	662.1 ± 148.4	7.4 ± 0.3	10.6 ± 4.6
G6	38.9 ± 14.8	3.9 ± 8.8	0.3 ± 0.3	5.7 ± 7.2	7.3 ± 3	720 ± 195.6	6.8 ± 0.4	10.9 ± 4.5
H1	50.3 ± 15	2.3 ± 3.2	0.1 ± 0.1	0.5 ± 0.6	7.5 ± 3.6	733.3 ± 131.4	7.6 ± 0.4	9.7 ± 4.7
H2	148.5 ± 104	35 ± 21.7	0.3 ± 0.4	9.1 ± 8.9	3.7 ± 2.2	1192.6 ± 523.6	7.4 ± 0.4	12.3 ± 4.8
H3	905 ± 465.1	6.1 ± 1.9	0.1 ± 0.1	3.4 ± 3.2	3 ± 1.4	3840.7 ± 2019.2	7.5 ± 0.4	11.0 ± 4.8
H4	89.7 ± 58	29.4 ± 19.8	0.1 ± 0.1	7.8 ± 9	2.3 ± 2.3	1032.7 ± 316.3	7.5 ± 0.3	14.1 ± 5.4
H5	25.5 ± 3.6	0.6 ± 1	0.1 ± 0	1.4 ± 3.3	6.6 ± 1.9	506.7 ± 53.9	7.4 ± 0.3	12.0 ± 4.1
H6	72.8 ± 35.1	7.4 ± 7.6	0.1 ± 0	2.9 ± 3.3	5 ± 2.5	594.9 ± 164.4	7.3 ± 0.5	11.1 ± 4.2
Average	**472.5 ± 1948.9**	**4.1 ± 10.7**	**0.2 ± 0.4**	**2.1 ± 4.5**	**7.0 ± 3.1**	**1556.7 ± 3593.3**	**7.3 ± 0.5**	**11.4 ± 4.9**

Seasonal nitrate concentrations (Tables 7-3 and 7-4) of the thirty sampling points varied widely (variable mean NO_3^- concentrations and relatively large standard deviations) between and among sampling points.

Table 7-3: Summary of NO_3^- concentrations and isotope statistics for 30 sampling points in winter (W).

Sampling code	NO_3^- (mg N L^{-1})			$\delta^{15}N$-NO_3^- (‰)			$\delta^{18}O$-NO_3^-(‰)		
	Min.	**Mean ± SD**	**Max.**	**Min.**	**Mean ± SD**	**Max.**	**Min.**	**Mean ± SD**	**Max.**
A1	1.7	6.1 ± 3.0	9.4	4.0	14.9 ± 5.3	21.7	2.1	10.8 ± 5.9	25.2
A2	5.9	18.8 ± 9.2	39.4	9.1	14.4 ± 4.3	23.4	3.7	7.8 ± 6.0	24.4
A3	0.5	5.2 ± 3.2	9.3	6.0	10.7 ± 2.1	14.2	3.2	9.4 ± 4.9	19.8
A4	0.1	2.7 ± 1.9	5.6	8.4	10.9 ± 1.8	14.0	1.0	8.0 ± 5.0	18.7
A5	2.9	7.0 ± 2.2	9.7	9.8	13.3 ± 1.9	16.1	7.5	11.7 ± 4.4	21.9
A6	0.2	4.6 ± 2.9	8.9	4.4	20.0 ± 6.6	28.0	-3.0	12.9 ± 6.8	23.8
AV. Class A	**0.1**	**7.1 ± 6.5**	**39.4**	**4.0**	**13.9 ± 5.0**	**28.0**	**-3.0**	**10.1 ± 5.6**	**25.2**
AGC1	0.2	1.9 ± 1.6	5.6	5.3	9.1 ± 2.4	12.8	1.4	8.2 ± 5.9	20.1
AGC2	8.8	13.0 ± 2.4	16.0	4.5	8.3 ± 1.7	11.4	-1.1	4.4 ± 3.3	10.9
AGC3	1.0	1.6 ± 0.4	2.6	4.8	10.5 ± 3.0	14.5	1.0	4.9 ± 2.3	9.6
AGC4	4.6	6.2 ± 1.0	7.6	5.7	9.7 ± 2.1	12.1	-6.7	6.5 ± 4.8	11.0
AGC5	7.5	10.8 ± 3.8	20.1	5.4	8.8 ± 2.0	11.6	2.4	7.1 ± 4.2	15.7
AGC6	6.6	11.4 ± 1.9	13.2	4.2	6.4 ± 1.5	8.5	2.1	5.3 ± 3.5	13.6
AV. Class AGC	**0.2**	**7.5 ± 5.0**	**20.1**	**4.2**	**8.8 ± 2.5**	**14.5**	**-6.7**	**6.1 ± 4.2**	**20.1**
AH1	0.1	22.3 ± 15.4	43.5	0.6	8.7 ± 4.8	17.1	-1.4	3.0 ± 3.0	8.1
AH3	8.0	18.7 ± 5.6	28.3	7.6	13.5 ± 3.3	18.2	0.5	9.0 ± 4.8	17.8
AH4	8.6	15.8 ± 4.2	23.8	9.6	13.6 ± 2.5	17.3	-1.2	8.9 ± 6.1	23.7
AH5	12.0	21.8 ± 7.4	34	12.9	15.4 ± 2.4	21.4	5.0	8.6 ± 2.5	13.0
AH6	10.3	18.0 ± 3.9	22.9	10.5	14.1 ± 2.2	18.4	5.4	9.0 ± 4.9	22.9
AV. Class AH	**0.1**	**18.6 ± 8.6**	**43.5**	**0.6**	**13.1 ± 3.7**	**21.4**	**-1.4**	**7.8 ± 4.9**	**23.7**
G1	6.4	19.7 ± 13.0	58.3	4.7	9.5 ± 3.3	15.6	8.9	14.4 ± 4.4	20.0
G2	1.3	17.6 ± 9.2	32.7	4.2	9.3 ± 3.7	17.6	12.3	18.0 ± 4.8	25.8
G3	0.1	3.4 ± 2.2	7.0	6.2	12.5 ± 3.1	16.6	2.3	10.9 ± 5.0	23.1
G4	0.1	4.9 ± 2.9	8.8	6.8	12.2 ± 4.0	19.3	5.0	10.2 ± 4.3	20.7
G5	1.7	4.6 ± 1.2	6.1	3.5	7.2 ± 2.9	13.2	9.2	19.4 ± 5.2	28.6
G6	11	26.3 ± 14.9	67.7	3.7	7.9 ± 2.2	10.5	9.0	16.2 ± 6.7	28.5
AV. Class G	**0.1**	**12.9 ± 12.5**	**67.7**	**3.5**	**9.7 ± 3.7**	**19.3**	**2.3**	**14.9 ± 6.0**	**28.6**
H1	2.3	4.2 ± 1.1	5.7	7.5	11.1 ± 2.7	14.7	-0.6	6.5 ± 4.4	13.6
H2	0.1	5.3 ± 4.6	12.0	8.7	14.8 ± 6.1	26.4	2.0	7.0 ± 2.9	13.3
H3	0.1	0.6 ± 0.8	2.6	-4.6	4.1 ± 7.2	11.8	-0.3	14.7 ± 11.6	32.7
H4	0.1	0.8 ± 0.9	2.4	-5.4	11.2 ± 8.7	19.6	-0.5	6.6 ± 4.4	11.3
H5	0.1	0.4 ± 0.2	0.7	-0.4	4.0 ± 2.5	7.9	-3.5	2.2 ± 3.1	7.6
H6	0.1	1.2 ± 1.1	3.0	-7.6	8.2 ± 6.2	17.3	2.4	6.5 ± 4.1	16.2
AV. Class H	**0.1**	**2.0 ± 2.7**	**12.0**	**-7.6**	**9.0 ± 6.6**	**26.4**	**-3.5**	**6.8 ± 6.2**	**32.7**

AV represents average

Obviously seasonal NO_3^- concentration variations can be found in class G (mean NO_3^- concentration in class G is 12.9 mg NL^{-1} in winter in Table 7-3 and 21.8 mg NL^{-1} in summer in Table 7-4), class A (mean NO_3^- concentration in class A is 7.1 mg NL^{-1} in winter in Table 7-3 and 1.6 mg NL^{-1} in summer in Table 7-4) and class AH (mean NO_3^- concentration in class

is AH is 18.6 mg NL^{-1} in winter in Table 7-3 and 10.4 mg NL^{-1} in summer in Table 7-4). The mean NO_3^- concentration in class AGC is almost the same in winter (7.5 mg NL^{-1}) as in summer (7.3 mg NL^{-1}) in Tables 7-3 and 7-4. Seasonal variation in class AGC is not expected due to ground water compensation. The mean NO_3^- concentration in class H is quite low compared to other classes, only 2.0 mg NL^{-1} in winter in Table 7-3 and 0.8 mg NL^{-1} in summer in Table 7-4. However, relatively higher NH_4^+ values were found.

Table 7-4: Summary of NO_3^- concentrations and isotope statistics for 30 sampling points in summer (S).

Sampling code	NO_3^- (mg N L^{-1})			$\delta^{15}N$-NO_3^-(‰)			$\delta^{18}O$-NO_3^-(‰)		
	Min.	Mean ± SD	Max.	Min.	Mean ± SD	Max.	Min.	Mean ± SD	Max.
A1	0.6	1.7 ± 1.1	3.4	15.0	18.1 ± 2.7	22.7	2.8	10.2 ± 3.8	-
A2	0.1	4.4 ± 4.7	12.0	6.9	16.1 ± 6.4	27.4	-6.0	18.9 ± 19.3	60.5
A3	0.1	0.9 ± 1.8	6.4	-10.3	3.1 ± 8.8	12.6	2.5	8.5 ± 4.3	15.3
A4	0.1	0.4 ± 0.6	1.8	-	-	-	-	-	-
A5	0.1	1.4 ± 2.6	8.2	-5.8	10.8 ± 8.6	17.4	-1.1	7.4 ± 5.4	11.6
A6	0.1	0.1 ± 0.1	0.2	6.5	7.7 ± 1.7	8.9	-0.9	10.2 ± 15.7	21.3
AV. Class A	**0.1**	**1.6 ± 2.9**	**12.0**	**-5.8**	**12.5 ± 8.2**	**27.4**	**-6**	**11.8 ± 11.7**	**60.5**
AGC1	0.1	0.2 ± 0.2	0.7	-3.7	3.1 ± 3.4	6.5	-11.7	3.9 ± 9.0	14.6
AGC2	6.4	15.1 ± 3.0	18.0	5.2	7.8 ± 1.3	9.6	0	6.5 ± 5.6	17.8
AGC3	0.6	1.2 ± 0.4	1.9	8.9	11.5 ± 2.1	16.2	-0.5	7.3 ± 3.5	13.9
AGC4	0.1	4.6 ± 2.0	6.7	7.6	12.2 ± 2.6	15.5	6.3	11.4 ± 4.3	20.1
AGC5	3.8	10.5 ± 2.9	15.0	6.6	10.2 ± 2.9	14.5	-1.4	12.0 ± 9.3	26.7
AGC6	9.9	12.0 ± 1.5	14.5	2.6	5.9 ± 2.1	8.8	0.4	10.1 ± 8.6	28.8
AV. Class AGC	**0.1**	**7.3 ± 6.0**	**18.0**	**-3.7**	**8.8 ± 3.8**	**16.2**	**-11.7**	**8.8 ± 7.3**	**28.8**
AH1	0.1	9.8 ± 9.1	19.6	9.9	11.8 ± 1.8	14.1	3.4	8.8 ± 5.3	15.3
AH3	2.1	11.6 ± 7.3	23.1	12.7	16.7 ± 3.5	22.1	6.6	14.5 ± 7.5	30.8
AH4	0.1	7.8 ± 4.8	12.8	13.6	16.2 ± 4.1	25.9	6.3	13.5 ± 6.0	22.7
AH5	2.5	11.4 ± 5.3	18.0	13.7	18.0 ± 2.3	21.8	3.1	13.4 ± 5.2	23.4
AH6	4.2	12.6 ± 6.3	23.2	10.3	18.0 ± 3.7	21.6	10	11.7 ± 1.7	15.3
AV. Class AH	**0.1**	**10.4 ± 6.4**	**23.2**	**9.9**	**16.7 ± 3.6**	**25.9**	**3.1**	**12.9 ± 5.5**	**30.8**
G1	16.0	45.8 ± 24.7	80.1	5.5	10.1 ± 3.6	19.0	17.4	26.6 ± 6.8	38.3
G2	9.7	25.9 ± 19.8	78.7	4.4	9.0 ± 3.0	13.7	10.2	32.0 ± 11.2	49.5
G3	1.1	5.9 ± 5.7	22.1	-2.3	5.1 ± 4.6	-	9.0	20.9 ± 6.8	-
G4	0.2	12.2 ± 20.1	71.6	3.7	10.8 ± 5.9	20.7	8.0	21.7 ± 8.3	39.1
G5	0.7	11.1 ± 14.9	51.3	-1.3	3.0 ± 2.6	5.7	22	24.9 ± 4.4	37.3
G6	5.1	31.0 ± 26.3	91.6	4.0	8.3 ± 4.2	18.3	14.0	27.2 ± 12.1	51.7
AV. Class G	**0.2**	**21.8 ± 23.5**	**91.6**	**-2.3**	**7.7 ± 4.8**	**20.7**	**8.0**	**25.6 ± 9.2**	**51.7**
H1	0.4	3.1 ± 4.1	13.0	8.5	13.1 ± 4.5	20.3	3.0	19.1 ± 11.5	29.2
H2	0.1	0.8 ± 1.4	4.5	10.7	16.0 ± 3.5	19.9	4.7	14.4 ± 9.0	24
H3	0.1	0.4 ± 0.4	1.2	-9.5	0.9 ± 5.2	6.9	-8.1	10.9 ± 11.9	32.8
H4	0.1	0.5 ± 0.7	1.7	4.7	12.9 ± 7.4	19.0	5.8	11.4 ± 7.3	19.7
H5	0.2	0.4 ± 0.1	0.7	-4.5	1.7 ± 2.6	4.7	-3.9	2.0 ± 8.3	26.8
H6	0.1	0.2 ± 0.1	0.4	1.9	6.9 ± 4.0	12.7	-3.3	7.7 ± 9.2	19.7
AV. Class H	**0.1**	**0.8 ± 1.8**	**13.0**	**-9.5**	**6.7 ± 7.1**	**20.3**	**-8.1**	**9.4 ± 10.9**	**32.8**

AV represents average

7.4.2 Multiple isotope approach ($\delta^{15}N$- and $\delta^{18}O$-NO_3^- and $\delta^{11}B$) for NO_3^- source identification

The total number of measured $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values is 561 and these values are summarized for winter and summer for individual sampling points in Tables 7-3 and 7-4. The seasonal mean $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of the 30 sampling points (24 months) for the five classes also varied differently. To derive qualitative information for identifying the predominant NO_3^- sources in the five different classes, a classical dual isotope bi-plot approach ($\delta^{15}N$- NO_3^- vs. $\delta^{18}O$-NO_3^-) could be applied in Figure 7-2.

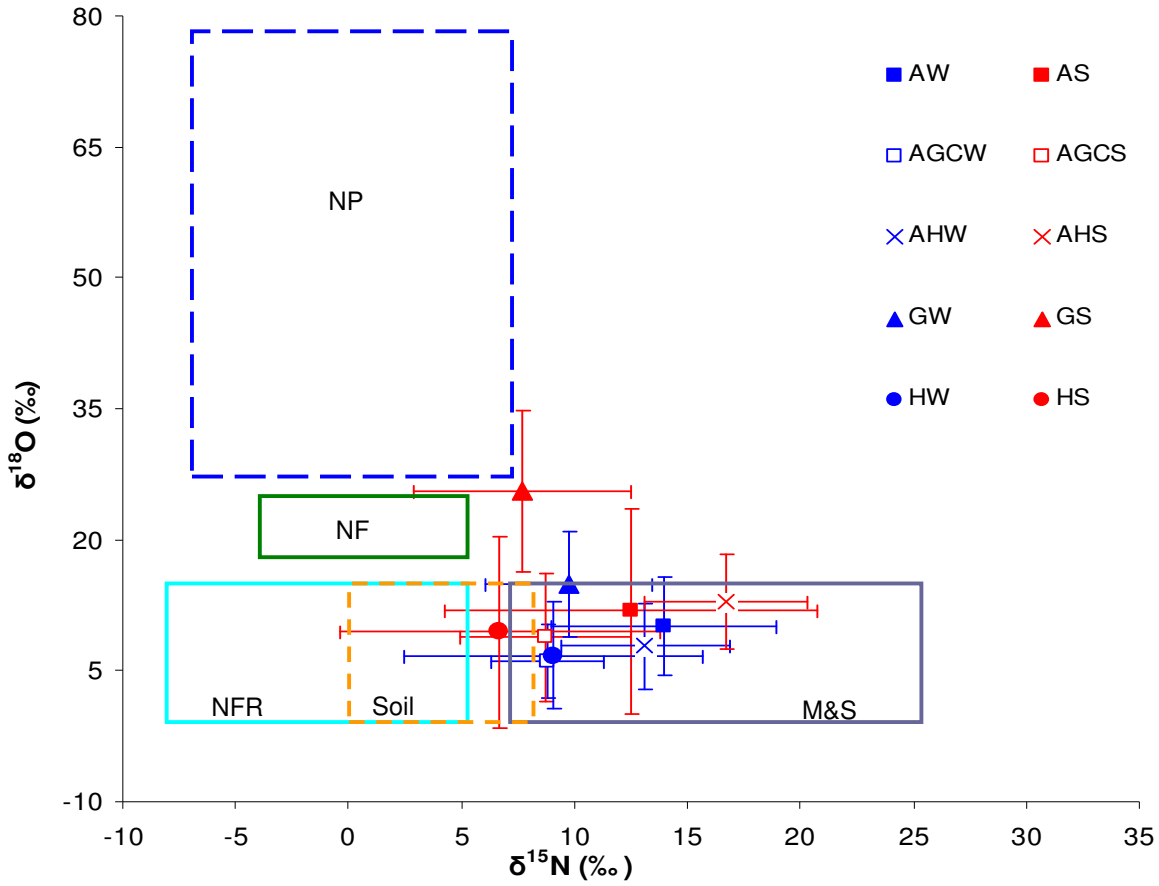


Figure 7-2: Seasonal average and standard deviation of $\delta^{15}N$ and $\delta^{18}O$-NO_3^- of five NO_3^- expert classes. Ranges of isotopic composition for five potential NO_3^- sources are determined by Xue et al. (2009) and indicated by boxes: NO_3^- in precipitation (NP), NO_3^- fertilizer (NF), NH_4^+ in fertilizer and rain (NFR), soil N (Soil) and manure and sewage (M&S). W represents winter and S represents summer. A: agriculture; AGC: agriculture with groundwater compensation; AH: a combination of agriculture with horticulture; G: greenhouses in an agricultural area; and H: households.

It is clear that the mean $\delta^{15}N$ and $\delta^{18}O\text{-}NO_3^-$ values of classes A and AH were relatively close and in the range of manure and sewage source. The distribution of classes AGC and H was in the window margin of manure and sewage source but shifted towards soil N. Only slight seasonal variations can be observed for all above classes. However, a more pronounced seasonal shift of $\delta^{15}N$ and $\delta^{18}O\text{-}NO_3^-$ was observed in class G, from NO_3^- in precipitation and NO_3^- fertilizer in summer, to manure and sewage in winter. In general, the manure and sewage source contributed to all five classes. However, it is impossible to specify which of both, as manure and sewage cannot be separated based on $\delta^{15}N$ and $\delta^{18}O$ values in Figure 7-2.

The $\delta^{11}B$ measurements can provide additional information for NO_3^- source identification for these five classes, as boron is widely present as a minor element in mineral N fertilizers in agriculture and as a bleaching agent for cleaning products in households. Figure 7-3 demonstrates the relationship between $\delta^{11}B$ values of 12 monthly composite samples for each of the five NO_3^- classes and B concentrations from October 2008 to September 2009. It is obvious that two groups can be identified: (1) classes A, AGC and AH with high $\delta^{11}B$ values, and (2) classes G and H with low $\delta^{11}B$ values. We found the majority of $\delta^{11}B$ measurements in classes A, AGC and AH were higher than 10‰, which was in the range of literature data for manure (Komor, 1997; Widory et al., 2004; Chetelat and Gaillardet 2005).

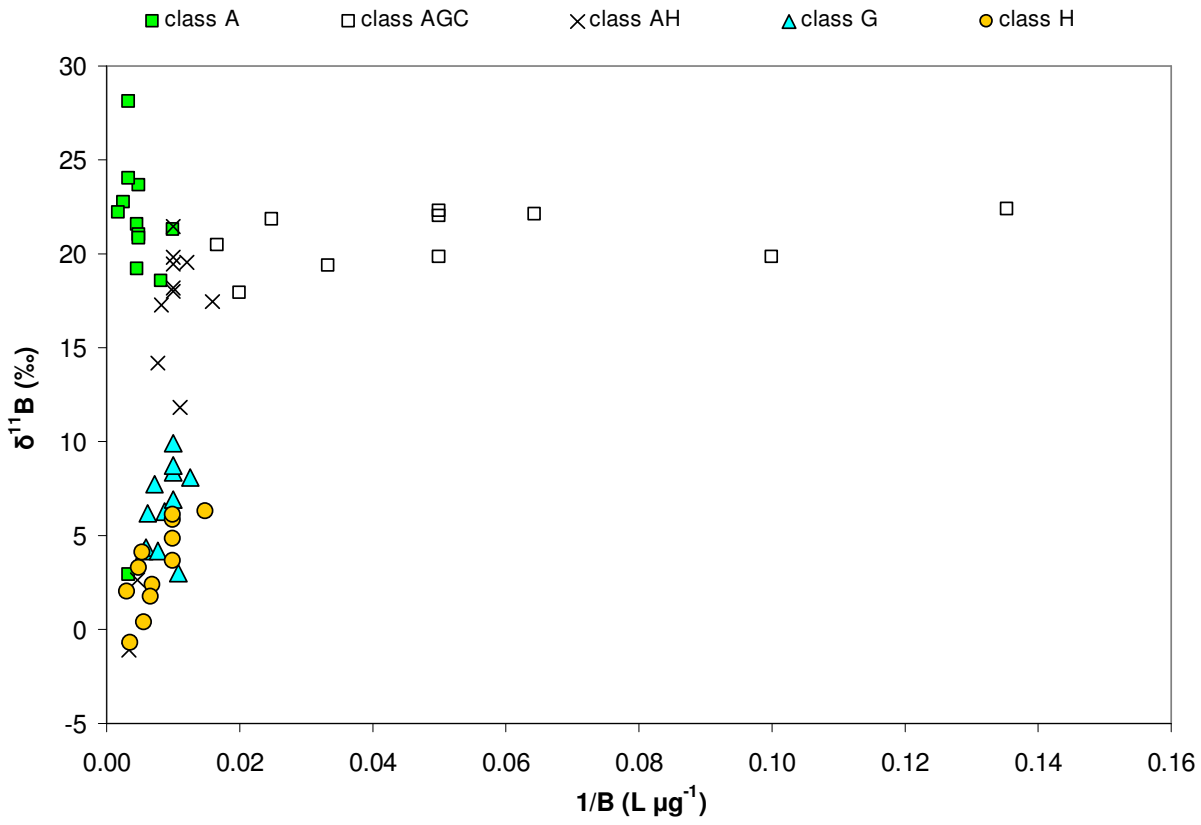


Figure 7-3: $\delta^{11}B$ vs. 1/B diagram for five NO_3^- classes from October 2008 to September 2009.

However, there was one outlier in class A and two outliers in class AH with comparatively low $\delta^{11}B$ values (lower than 5‰), which could be due to mineral N input in summer time in agricultural areas. B concentrations ranged from 100 to 550 μg L^{-1} for class A and from 8 to 60 μg L^{-1} for class AGC and from 60 to 300 μg L^{-1} for class AH. Furthermore, the variability in 1/B and relative stability in $\delta^{11}B$ indicated that dilution with clean ground water occured in class AGC. Sampling points for class G were also located in agricultural` areas, but with intensive greenhouse activity and showed $\delta^{11}B$ values (between 3.0 and 9.9‰) within the literature range for mineral fertilizer (Komor, 1997; Widory et al., 2004; Chetelat and Gaillardet 2005), indicating mineral N input as a dominant source. Sampling points for class H are located in an area dominated by households and the $\delta^{11}B$ values for class H varied between -0.7 and 6.3‰, occurring in the literature range of sewage (Vengosh et al., 1994; Bassett et al., 1995; Leenhouts et al., 1998; Vengosh et al., 1999; Widory et al., 2004; Widory et al., 2005; Seiler, 2005). Mineral fertilizers as the dominant source can be excluded as the mean $\delta^{15}N\text{-}NO_3^-$ values of class H pointed towards manure or sewage. Therefore, the dominant source in class H was sewage. Boron concentrations varied from 80 to 180 μg L^{-1} for class G and from 70 to 330 μg L^{-1} for class H.

Thus, application of a multiple isotope approach ($\delta^{15}N$-, $\delta^{18}O\text{-}NO_3^-$, and $\delta^{11}B$) can provide more accurate information for NO_3^- source identification. Manure was the dominant source for class A, class AGC and class AH both in winter and summer, while sewage was the dominant source for class H. Although $\delta^{11}B$ and $\delta^{15}N$ data indicated mineral fertilizer could be the dominant source for class G, we cannot exclude considerable contribution from NO_3^- in precipitation in summer and manure in winter, which was evidenced from the classical dual isotope bi-plot approach (Figure 7-2). In addition, B content in fertilizers could range from below detection limit for some brands of ammonium nitrate fertilizer up to 382 mg kg^{-1} in magnesium sulfate (Komor, 1997; Tirez et al., 2010). Thus, we cannot exclude an influence from other non-N fertilizers. Qualitative assessment of predominant NO_3^- sources was obtained via a classical dual isotope bi-plot approach ($\delta^{15}N\text{-}NO_3^-$ vs. $\delta^{18}O\text{-}NO_3^-$) and $\delta^{11}B$ and B concentration analysis. However, a more quantitative assessment of NO_3^- source apportionment should be developed via SIAR.

7.4.3 Application of SIAR for estimating multiple NO_3^- source contribution

SIAR was applied to estimate five potential NO_3^- source apportionments (NO_3^- in precipitation (NP), NO_3^- fertilizer (NF), NH_4^+ in fertilizer and rain (NFR), soil N and manure

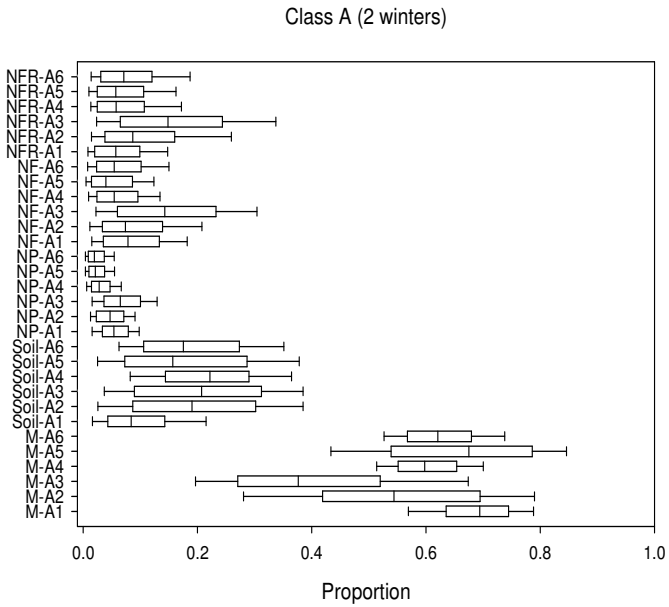
or sewage (M or S)) in the five different classes (A, AGC, AH, G and H) in different seasons (two winters and two summers) and the complete monitoring period based on 561 $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of the 30 sampling points. The isotope values of the five sources were assumed to fall into the source boxes established from literature data (Figure 7-2). The SIAR mixing model outputs revealed a high variability in contribution of the five potential NO_3^- sources. The contribution ranges of each NO_3^- source to the 30 sampling points are shown in Figure 7-4.

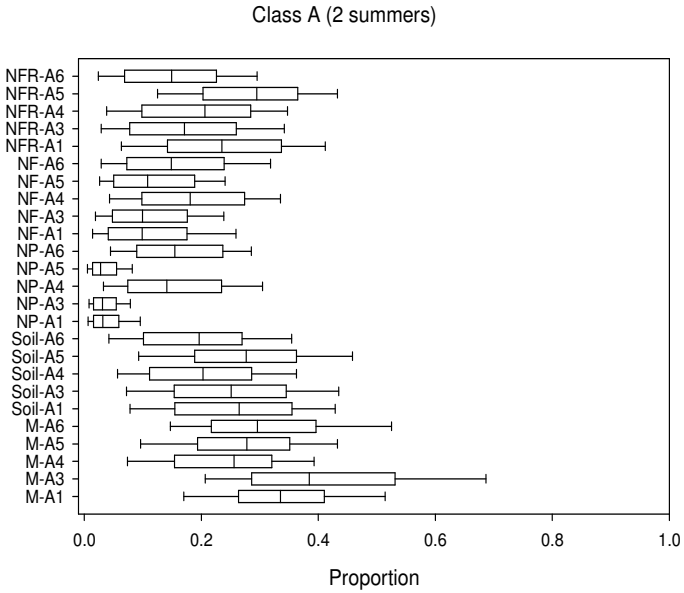
For the six sampling points in class A (Figure 7-4a), manure contributed about 60% in winter, and the remaining four sources contributed 40% in total. In summer, the proportional contribution of manure decreased to 30% and the other sources relatively increased (mean values around 10 to 20%). The SIAR output for the entire two year data set demonstrated a similar source contribution pattern as in winter. Thus, winter NO_3^- contributions has relatively more effect than summer NO_3^- contributions.

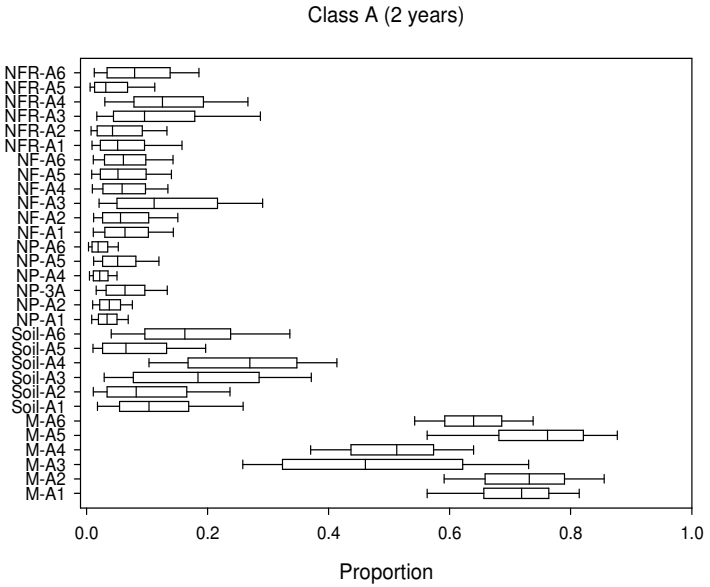
There was no apparent seasonal variation for the five NO_3^- sources during winter, summer and the entire two years for class AGC, except for the relatively low manure source contribution for one sampling point in summer (Figure 7-4b). Manure always tended to be the dominant source (mean values between 50% and 60%), followed by soil N (mean values between 20% and 30%). The other three sources contributed ca. 10%. This lack of variation could be a result of attenuation by clear ground water in these areas.

The source contributions in class AH (Figure 7-4c) showed a similar seasonal pattern as in class A. Manure was the dominant source in winter (mean value ca. 60%) and for the entire two years (mean value ca. 70%), and NO_3^- in precipitation contributed least (mean value < 5%). The other three sources were intermediate (mean values higher than 5% and lower than 10%). In summer, the proportional contribution of manure decreased to ca. 40% and the other sources relatively increased (mean values around 10 to 20%).

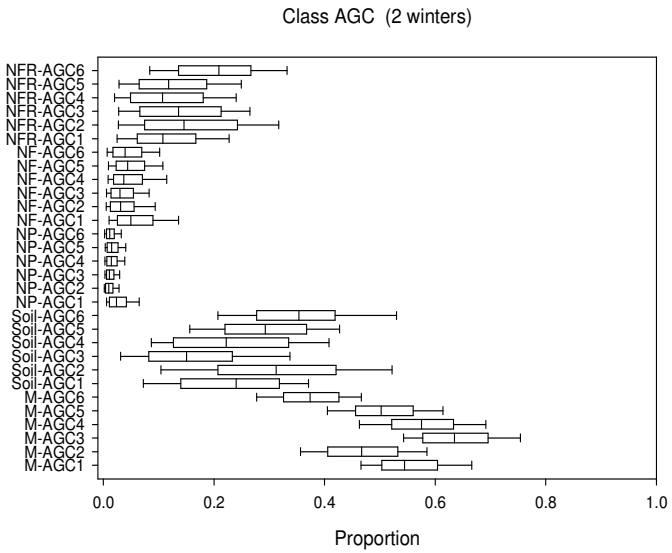
(a)

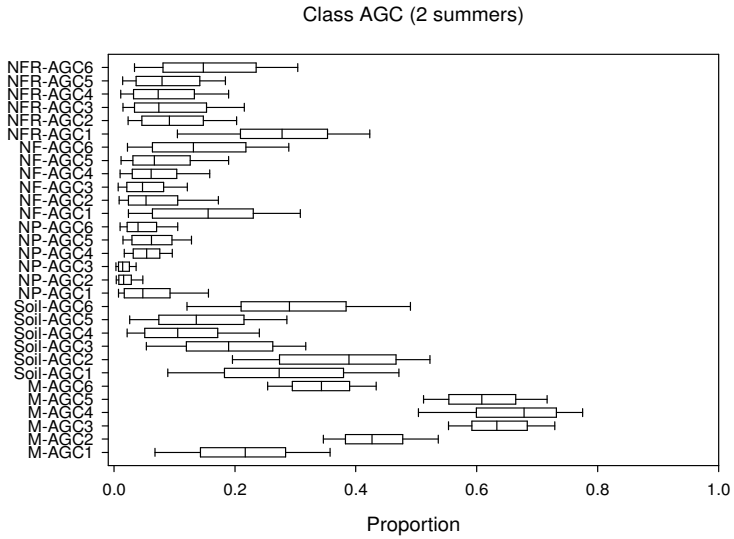

Class A (2 winters)
NFR-A6
NFR-A5
NFR-A4
NFR-A3
NFR-A2
NFR-A1
NF-A6
NF-A5
NF-A4
NF-A3
NF-A2
NF-A1
NP-A6
NP-A5
NP-A4
NP-A3
NP-A2
NP-A1
Soil-A6
Soil-A5
Soil-A4
Soil-A3
Soil-A2
Soil-A1
M-A6
M-A5
M-A4
M-A3
M-A2
M-A1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion

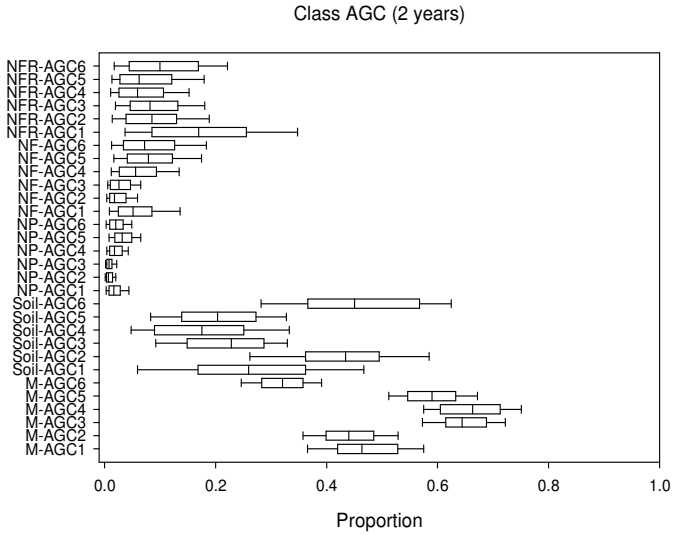

Class A (2 summers)
NFR-A6
NFR-A5
NFR-A4
NFR-A3
NFR-A1
NF-A6
NF-A5
NF-A4
NF-A3
NF-A1
NP-A6
NP-A5
NP-A4
NP-A3
NP-A1
Soil-A6
Soil-A5
Soil-A4
Soil-A3
Soil-A1
M-A6
M-A5
M-A4
M-A3
M-A1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion


Class A (2 years)
NFR-A6
NFR-A5
NFR-A4
NFR-A3
NFR-A2
NFR-A1
NF-A6
NF-A5
NF-A4
NF-A3
NF-A2
NF-A1
NP-A6
NP-A5
NP-A4
NP-3A
NP-A2
NP-A1
Soil-A6
Soil-A5
Soil-A4
Soil-A3
Soil-A2
Soil-A1
M-A6
M-A5
M-A4
M-A3
M-A2
M-A1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion

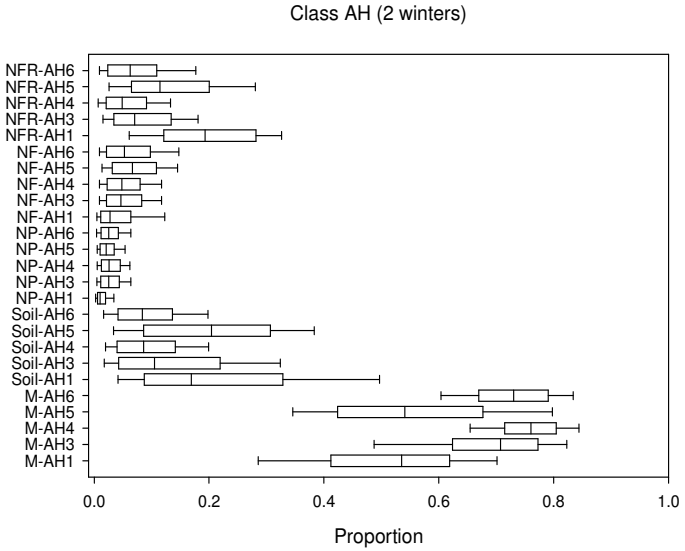
(b)

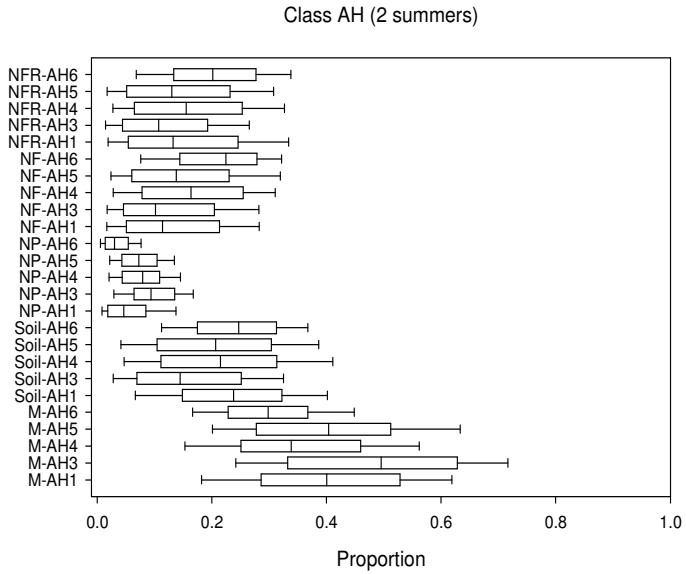

Class AGC (2 winters)
NFR-AGC6
NFR-AGC5
NFR-AGC4
NFR-AGC3
NFR-AGC2
NFR-AGC1
NF-AGC6
NF-AGC5
NF-AGC4
NF-AGC3
NF-AGC2
NF-AGC1
NP-AGC6
NP-AGC5
NP-AGC4
NP-AGC3
NP-AGC2
NP-AGC1
Soil-AGC6
Soil-AGC5
Soil-AGC4
Soil-AGC3
Soil-AGC2
Soil-AGC1
M-AGC6
M-AGC5
M-AGC4
M-AGC3
M-AGC2
M-AGC1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion

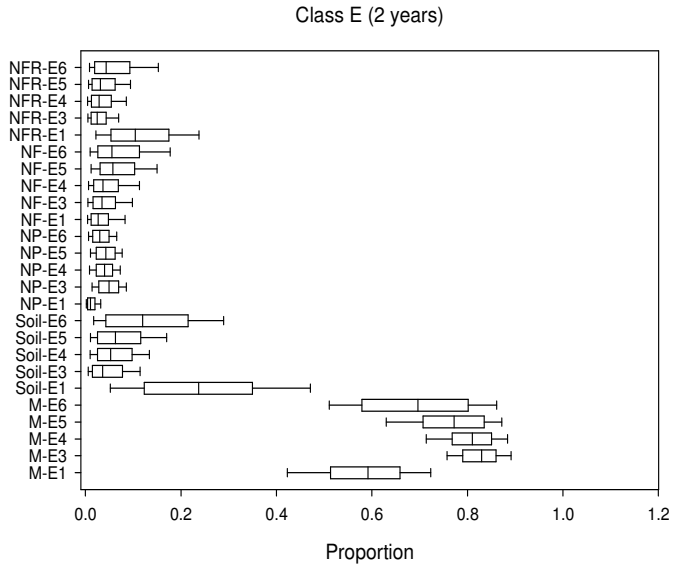

Class AGC (2 summers)
NFR-AGC6
NFR-AGC5
NFR-AGC4
NFR-AGC3
NFR-AGC2
NFR-AGC1
NF-AGC6
NF-AGC5
NF-AGC4
NF-AGC3
NF-AGC2
NF-AGC1
NP-AGC6
NP-AGC5
NP-AGC4
NP-AGC3
NP-AGC2
NP-AGC1
Soil-AGC6
Soil-AGC5
Soil-AGC4
Soil-AGC3
Soil-AGC2
Soil-AGC1
M-AGC6
M-AGC5
M-AGC4
M-AGC3
M-AGC2
M-AGC1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion


Class AGC (2 years)
NFR-AGC6
NFR-AGC5
NFR-AGC4
NFR-AGC3
NFR-AGC2
NFR-AGC1
NF-AGC6
NF-AGC5
NF-AGC4
NF-AGC3
NF-AGC2
NF-AGC1
NP-AGC6
NP-AGC5
NP-AGC4
NP-AGC3
NP-AGC2
NP-AGC1
Soil-AGC6
Soil-AGC5
Soil-AGC4
Soil-AGC3
Soil-AGC2
Soil-AGC1
M-AGC6
M-AGC5
M-AGC4
M-AGC3
M-AGC2
M-AGC1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion

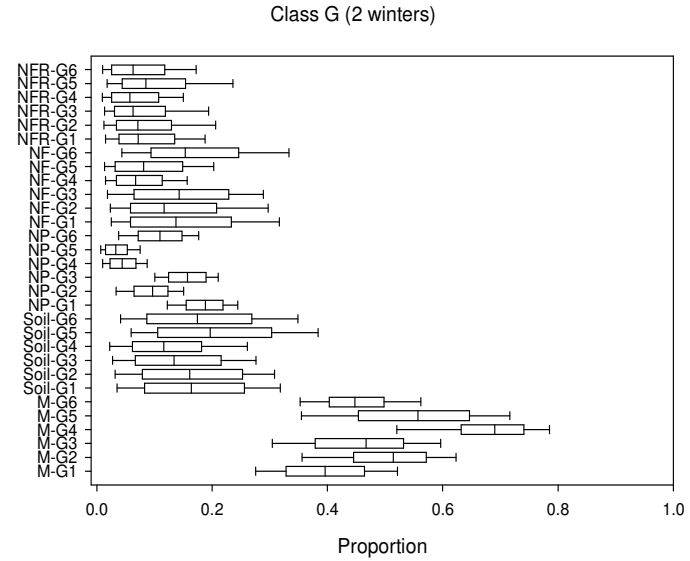
(c)

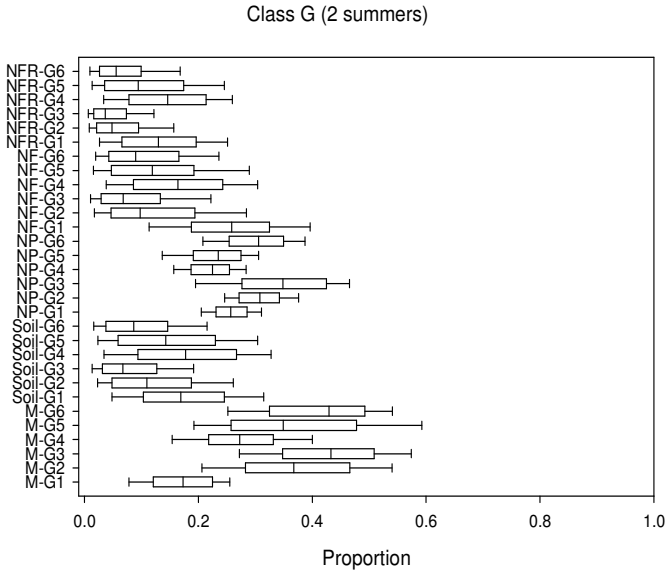

Class AH (2 winters)
NFR-AH6
NFR-AH5
NFR-AH4
NFR-AH3
NFR-AH1
NF-AH6
NF-AH5
NF-AH4
NF-AH3
NF-AH1
NP-AH6
NP-AH5
NP-AH4
NP-AH3
NP-AH1
Soil-AH6
Soil-AH5
Soil-AH4
Soil-AH3
Soil-AH1
M-AH6
M-AH5
M-AH4
M-AH3
M-AH1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion

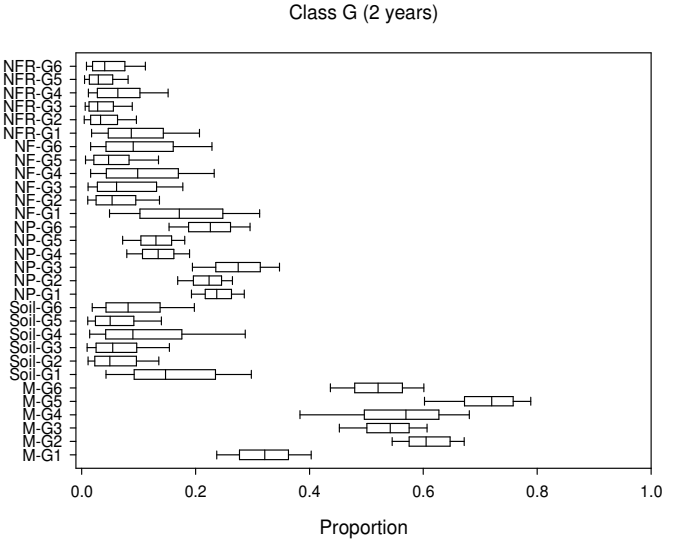

Class AH (2 summers)
NFR-AH6
NFR-AH5
NFR-AH4
NFR-AH3
NFR-AH1
NF-AH6
NF-AH5
NF-AH4
NF-AH3
NF-AH1
NP-AH6
NP-AH5
NP-AH4
NP-AH3
NP-AH1
Soil-AH6
Soil-AH5
Soil-AH4
Soil-AH3
Soil-AH1
M-AH6
M-AH5
M-AH4
M-AH3
M-AH1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion


Class E (2 years)
NFR-E6
NFR-E5
NFR-E4
NFR-E3
NFR-E1
NF-E6
NF-E5
NF-E4
NF-E3
NF-E1
NP-E6
NP-E5
NP-E4
NP-E3
NP-E1
Soil-E6
Soil-E5
Soil-E4
Soil-E3
Soil-E1
M-E6
M-E5
M-E4
M-E3
M-E1
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Proportion

(d)


Class G (2 winters)
NFR-G6
NFR-G5
NFR-G4
NFR-G3
NFR-G2
NFR-G1
NF-G6
NF-G5
NF-G4
NF-G3
NF-G2
NF-G1
NP-G6
NP-G5
NP-G4
NP-G3
NP-G2
NP-G1
Soil-G6
Soil-G5
Soil-G4
Soil-G3
Soil-G2
Soil-G1
M-G6
M-G5
M-G4
M-G3
M-G2
M-G1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion


Class G (2 summers)
NFR-G6
NFR-G5
NFR-G4
NFR-G3
NFR-G2
NFR-G1
NF-G6
NF-G5
NF-G4
NF-G3
NF-G2
NF-G1
NP-G6
NP-G5
NP-G4
NP-G3
NP-G2
NP-G1
Soil-G6
Soil-G5
Soil-G4
Soil-G3
Soil-G2
Soil-G1
M-G6
M-G5
M-G4
M-G3
M-G2
M-G1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion


Class G (2 years)
NFR-G6
NFR-G5
NFR-G4
NFR-G3
NFR-G2
NFR-G1
NF-G6
NF-G5
NF-G4
NF-G3
NF-G2
NF-G1
NP-G6
NP-G5
NP-G4
NP-G3
NP-G2
NP-G1
Soil-G6
Soil-G5
Soil-G4
Soil-G3
Soil-G2
Soil-G1
M-G6
M-G5
M-G4
M-G3
M-G2
M-G1
0.0
0.2
0.4
0.6
0.8
1.0
Proportion

(e)

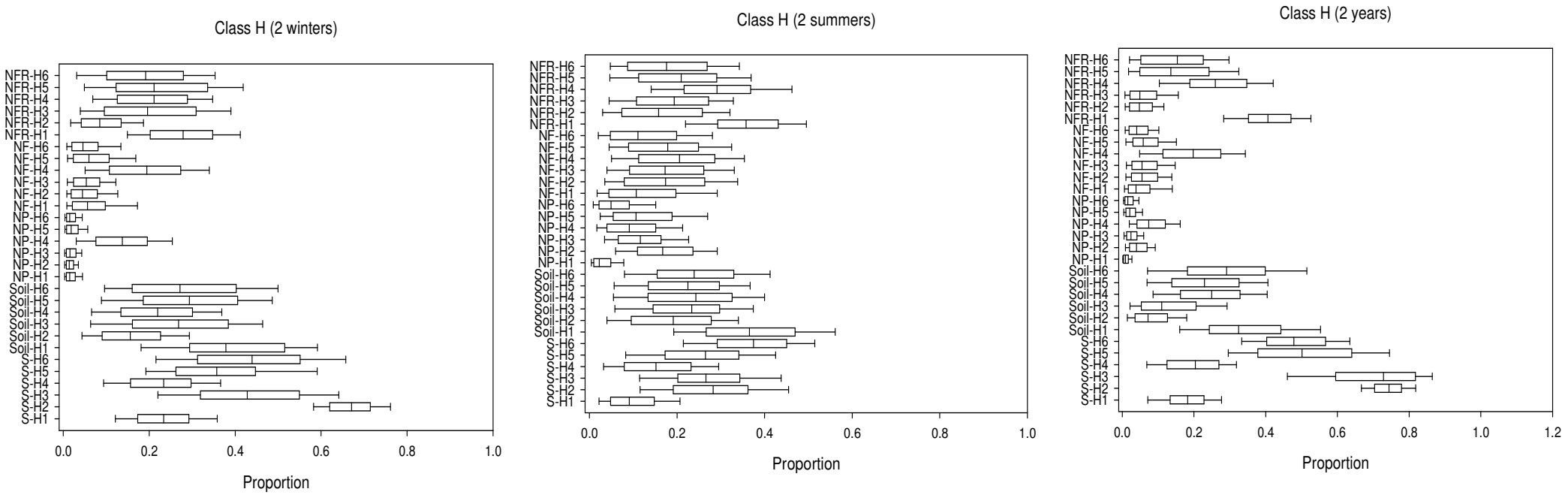


Figure 7-4: Contributions of five potential NO_3^- sources for five classes estimated by SIAR. M represents manure; S represents sewage; Soil represents soil N; NP represents NO_3^- in precipitation; NF represents NO_3^- fertilizer and NFR represents NH_4^+ in fertilizer and rain. Boxplots illustrate the 25th, 50th, and 75th percentiles; the whiskers indicate the 5th and 95th percentiles; abbreviations of 5 NO_3^- expert classes are given in Figure 7-2.

For the six sampling points in class G (Figure 7-4d), manure was the dominant source in winter (mean value ca. 50%), and the other four sources contributed more or less equally (mean values between 10 and 20%). In summer, the contribution of manure was lower (mean value ca. 30%) but NO_3^- in precipitation contribution was higher (mean value ca. 30%), with no clear variation in the other source contributors. The SIAR outputs for the entire two years revealed a similar source contribution pattern as in winter, indicating a great effect of winter contribution on the entire data set. Manure was the dominant source for class G throughout the sampling period, as there was agriculture in the vicinity of greenhouse and manure was the main fertilizer applied in Flanders (Eppinger et al., 2005). The high contribution from NO_3^- in precipitation (mean value ca. 30%) and high mean NO_3^- concentration (21.8 mg NL^{-1}) both occurred in summer. One possible explanation could be that farmers with greenhouses collected rainwater in summer to produce fertilizer solutions for crop growth. They irregularly discharge this waste water into surface water after a certain number of re-circulations. The released waste water might lead to high NO_3^- concentrations bearing a mineral fertilizer and rain water signal. Furthermore, the $\delta^{11}B$ values (between 3.0 and 9.9‰ in Figure 7-3) also indicated mineral N input as a dominant source.

Since the $\delta^{11}B$ values of class H (households)varied in the literature range of sewage, manure was not considered in this class. Therefore, the "M" output from SIAR was considered as sewage. Sewage contributed most (mean values between 40 and 50%) and NO_3^- in precipitation contributed least (mean values < 10%) in winter and for the entire two year data set in class H (Figure 7-4e). The other three sources contributed between 10 and 30%. In summer, all five sources contributed equally (mean values ca. 20%).

The SIAR outputs pointed out dominant and other important potential NO_3^- sources and also reflected temporal variations of the corresponding NO_3^- source contributions. In general, SIAR estimated "manure" and "sewage" as major, "NO_3^- in precipitation" as minor in winter, while the other three sources, "soil N", "NO_3^- fertilizer" and "NH_4^+ in fertilizer and rain" were intermediate for classes A, AGC, AH and H. For class G, manure was also the dominant source in winter, and the other four sources contributed more or less equally. In summer, the SIAR estimated manure as the dominant NO_3^- source for classes A and AH as well. However, the proportional contribution of manure decreased compared to winter, and the other sources relatively increased. The five sources showed a similar pattern in summer and in winter for class AGC. The source contributions of manure and NO_3^- in precipitation were dominant for class G, and the other three sources contributed equally. For class H, all the five sources contributed in a similar range in summer.

In Chapter 6, we showed that SIAR is a useful approach to estimate temporal and spatial variations of different NO_3^- sources. However, its resolution is largely determined by the temporal variability of the isotopic composition of NO_3^- in the mixture and the uncertainty on the isotopic composition of the different NO_3^- sources.

7.4.4 Retrieval of expert classification of NO_3^- sources

The "SIAR fingerprint" indicated that source contributions were variable within and between sampling points. Thus, some sampling points in the same expert class behaved quite differently and could have been classified into the wrong class only based on expert knowledge. Therefore, it is necessary to verify expert classification of the 30 sampling points based on the SIAR fingerprint using a k-mean clustering approach. Based on a multiple isotope approach ($\delta^{15}N\text{-}NO_3^-$, $\delta^{18}O\text{-}NO_3^-$ and $\delta^{11}B$) and SIAR outputs, "manure" was considered as the dominant source for classes A, AGC, AH and G and "sewage" was considered as the dominant source for class H. Therefore, we decided to separate households (6 sampling points) as a priori from the other four classes. Thus, we re-clustered the sampling points for the other four classes for winter and summer seasons separately. Furthermore, in Chapter 3, we explored a decision tree model to evaluate the CCI (percentage of correctly classified instances) for the 30 sampling points compared to the expet classification. Therefore, comparison of expert classification, decision tree model performance and k-means clustering results are shown in Table 7-5 for winter and in Table 7-6 for summer.

The number of sampling points in summer (22 sampling points) was less than in winter (23 sampling points) because sampling point A4 in class A was dry during summer. Sampling point AH2 was dry during the entire monitoring period. It is clear that the silhouette values were identical for 3 and 4 clusters, both 0.6 in winter and summer, indicating a reasonable clustering for both clusters. However, some sampling points defined in the same class by experts were clustered into different clusters. Considering the clustering results in Table 7-5 (winter), 2 sampling points in class A, 1 sampling point in class AGC, 5 sampling points in class AH and 2 sampling points in class G were clustered into different clusters, while the others were clustered into the same cluster as the expert classification. Therefore, it indicates that the dominant clusters correspond to classes A (cluster 2), AGC (cluster 3) and G (cluster 1) for the case of 3 clusters, as at least 4 out of 6 sampling points in each class remained in the same cluster. The sampling points in class AH were clustered into the classes A and AGC. In the case of 4 clusters in winter, it seems that the dominant clusters correspond to classes A

(cluter 2), AGC (cluster 3), AH (cluster 4) and G (cluster 1), but nearly half of the sampling points in one same expert class were clustered into different clusters: 3 sampling points in class A, 3 sampling points in class AGC, 2 sampling points in class AH and 2 sampling points in class G.

Table 7-5: Comparison of expert classification, decision tree model performance from Chapter 3 and k-means clustering results for the sampling points in expert classes A, AGC, AH and G for winter.

Expert classification	Sampling point	3 clusters via k-means	4 clusters via k-means	CCI based on expert classification by decision tree model (%)
A	A1	**cluster 3**	cluster 2	86
	A2	cluster 2	**cluster 4**	88
	A3	cluster 2	cluster 2	86
	A4	cluster 2	cluster 2	83
	A5	cluster 2	**cluster 4**	81
	A6	**cluster 1**	**cluster 1**	76
AGC	AGC1	cluster 3	**cluster 2**	85
	AGC2	cluster 3	cluster 3	97
	AGC3	**cluster 2**	**cluster 4**	86
	AGC4	cluster 3	**cluster 2**	93
	AGC5	cluster 3	cluster 3	93
	AGC6	cluster 3	cluster 3	91
AH	AH1	**cluster 3**	**cluster 2**	61
	AH3	**cluster 2**	cluster 4	56
	AH4	**cluster 2**	cluster 4	56
	AH5	**cluster 3**	**cluster 2**	65
	AH6	**cluster 2**	cluster 4	53
G	G1	cluster 1	cluster 1	84
	G2	cluster 1	cluster 1	89
	G3	**cluster 2**	**cluster 4**	76
	G4	**cluster 3**	**cluster 2**	83
	G5	cluster 1	cluster 1	80
	G6	cluster 1	cluster 1	62
Silhouette		0.6	0.6	

In Table 7-6 (summer), it seems that the dominant clusters correspond to classes A (cluster 2), AGC (cluster 3) and G (cluster 1) for both 3 clusters and 4 clusters. Most of the sampling points remained clustered into the classes defined by experts, but the sampling points in class AH were mainly clustered into class A (cluster 2). For 3 clusters, 3 sampling points in class AGC and 5 sampling point in class AH were clustered differently compared to expert classification. For 4 clusters, 2 sampling points in class A, 3 sampling points in class AGC, 5 sampling points in class AH and 2 sampling points in class G were clustered differently.

When looking into the CCI of the samling points using the decision tree model of Chapter 3 in Tables 7-5 and 7-6, it is obvious that the values were in a wide range from 53 to 97% for winter and from 47 to 94% for summer. The 5 sampling points with very low CCI in class AH (47-65%) were all clustered into either class A or class AGC, except for k-means clustering results for 4 clusters in Table 7-5. This information also indicated a high probability that the sampling points in class AH were misclassified by experts. Furthermore, the majority of the sampling points which were clustered into different clusters compared to its expert class, showed also relatively low CCI. However, some sampling points with high CCI cannot be excluded.

Table 7-6: Comparison of expert classification, decision tree model performance from Chapter 3 and k-means clustering results for the sampling points in expert classes A, AGC, AH and G for summer.

Expert classification	Sampling Point	3 clusters via k-means	4 clusters via k-means	CCI based on expert classification by decision tree model (%)
A	A1	cluster 2	cluster 2	91
	A2	cluster 2	**cluster 4**	84
	A3	cluster 2	cluster 2	87
	A5	cluster 2	cluster 2	75
	A6	cluster 2	**cluster 4**	82
AGC	AGC1	cluster **2**	**cluster 2**	81
	AGC2	cluster **2**	**cluster 2**	82
	AGC3	cluster 3	cluster 3	85
	AGC4	cluster 3	cluster 3	91
	AGC5	cluster 3	cluster 3	91
	AGC6	cluster **2**	**cluster 2**	94
AH	AH1	**cluster 2**	**cluster 2**	58
	AH3	**cluster 3**	**cluster 3**	61
	AH4	**cluster 2**	**cluster 2**	63
	AH5	**cluster 2**	**cluster 2**	48
	AH6	**cluster 2**	**cluster 2**	47
G	G1	cluster 1	cluster 1	77
	G2	cluster 1	cluster 1	82
	G3	cluster 1	**cluster 4**	76
	G4	cluster 1	cluster 1	95
	G5	cluster 1	**cluster 4**	72
	G6	cluster 1	cluster 1	67
Silhouette		0.6	0.6	

To estimate which number of cluster (3 clusters or 4 clusters) was closest to the expert classification, the “Rand index” for clustering comparison is shown in Table 7-7. The higher the index value is, the closer k-means clustering is to the expert classification. The values

were all similar 0.7. We suggest that 3 clusters is a reasonable choice, which showed the same silhouette value and Rand index value as 4 clusters. Furthermore, it keeps the majority of the sampling points from the same expert class clustered into the same cluster. Finally, we concluded that using an isotopic fingerprinting the sampling points were divided into four classes both in winter and summer: class A (agriculture), class AGC (agriculture with groundwater compensation), classG (greenhouses in an agricultural area), and class H (households).

Table 7-7: Comparison of expert classification and k-means clustering results in terms of the Rand index.

Season	Cluster comparison	Rand index
Winter	expert classification vs. 3 clusters	0.7
	expert classification vs. 4 clusters	0.7
Summer	expert classification vs. 3 clusters	0.7
	expert classification vs. 4 clusters	0.7

7.4.5 Final decision tree models for classification of NO_3^- sources

The 30 sampling points have been classified into 5 classes (classes A, AGC, AH, G and H) based on expert knowledge and 4 classes (classes A, AGC, G and H) based on the k-means clustering using the output of the SIAR model from two-year δ^{15}N- and δ^{18}O-NO_3^- data. Finally, we developed decision tree models using physico-chemical data alone and in combination with N and O isotopes from the monitoring period October 2007 to September 2009. For winter and summer, we compared the performance of both decision tree models with the NO_3^- source classification based on expert knowledge and the k-means clustering approach.

The seasonal CCI of different decision tree models for the 29 sampling points (sampling point 130350 was dry during the entire monitoring period) are shown in Table 7-8. It is clear that for expert classification and for classification using the k-means clustering approach, similar CCI's for winter (CCI between 68 and 70%) and summer (CCI between 74 and 75%) were obtained. This indicates that N and O isotope data could be used to develop a NO_3^- classification when expert knowledge is absent or not possible. However, the CCI of the decision tree models using δ^{15}N- and δ^{18}O-NO_3^- as attributes are similar to the ones without

using isotopes. The latter indicates that isotope data did not improve the decision tree model performance. This may result from the complex land use in Flanders, where multiple N sources were applied, resulting in NO_3^- isotope values that are largely scattered.

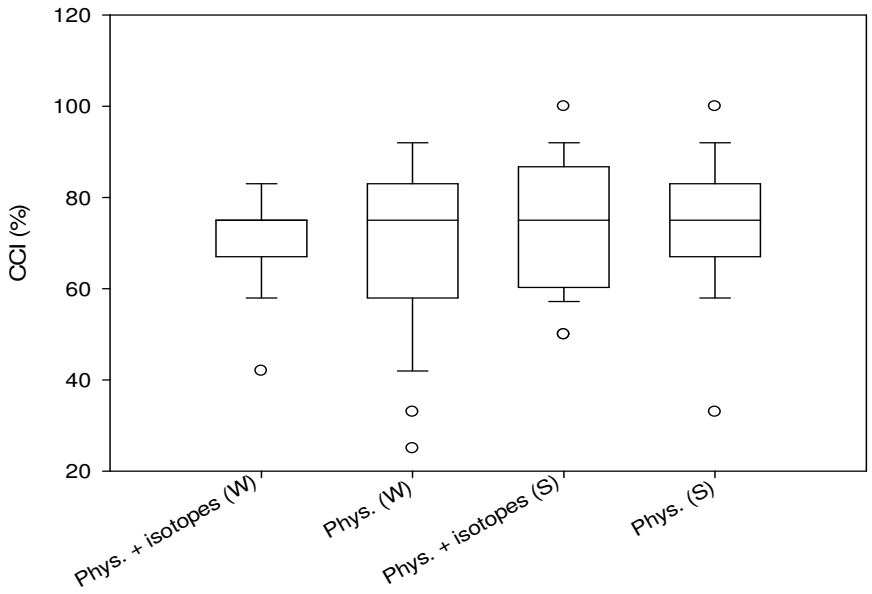
Table 7-8: The performance of decision tree models uing "Phys." and "Phys. + isotope" as attributes compared to expert classification and the k-means clustering approach.

Attributes for decision tree model development	**CCI based on expert classification (%)**		**CCI based on k-means clustering approach (%)**	
	Winter	Summer	Winter	Summer
Phys.	70	75	68	74
Phys. + isotope	70	75	70	74

Phys. + isotope represents both physicochemical properties and $\delta^{15}N$ and $\delta^{18}O\text{-}NO_3^-$ were used as attributes for decision tree model development; **Phys.** represents physicochemical properties were used as attributes for decision tree model development;

Information for performance of different decision tree models on individual sampling points can be found in Figure 7-5. The CCI of individual sampling points ranged from 25 to 100% compare to the expert classification (Figure 7-5a), which is much wider than the classification for the k-means clustering method (Figure 7-5b) (50 to 100%). The typical CCI (the 5th and 95th percentiles in the boxplot) were between 42 and 92% for the decision tree models comparing expert classification and between 50 and 92% for the decision tree models comparing the k-means clustering approach. The narrower range in the latter classification method implies that the decision tree models were relatively more stable when verifying the k-means clustering classification than expert knowledge classification. Furthermore, the outliers with low CCI in Figure 7-5a for the expert classification were all in the typical range of CCI in Figure 7-5b for the k-means clustering approach, which implies that the the k-means clustering approach improves the classification of sampling points with low CCI in expert classification.

(a) expert classification



(b) k-mean clustering

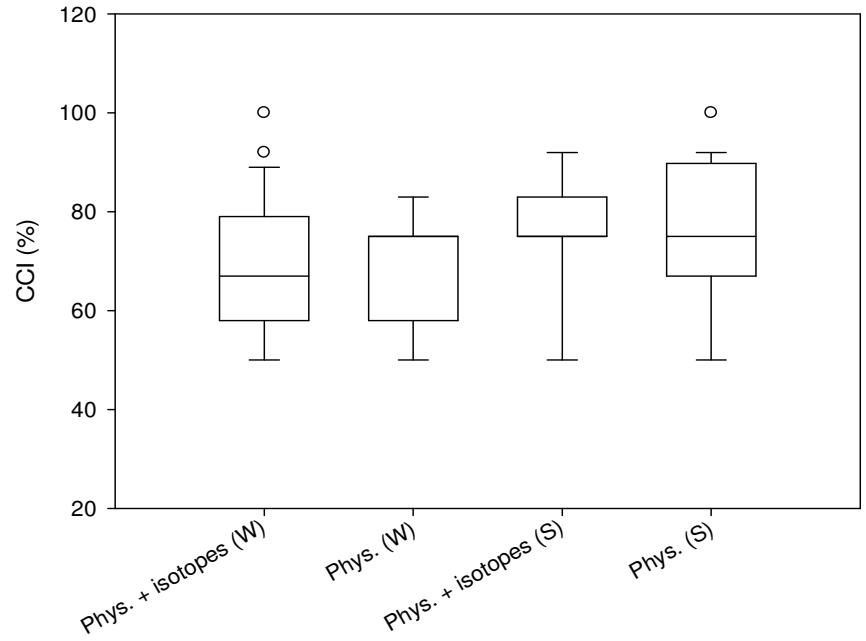


Figure 7-5: Boxplot of CCI (%) for the 29 sampling points obtained from different decision tree models uing "Phys." and "Phys. + isotope" as attributes compared to (a) expert classification and (b) k-means clustering. Phys.+isotope represents both physicochemical properties and $\delta^{15}N$ and $\delta^{18}O\text{-}NO_3^-$ were used as attributes for decision tree model development; Phys. represents physicochemical properties were used as attributes for decision tree model development; W represents winter; S represents summer. Boxplots illustrate the 25th, 50th and 75th percentiles; the whiskers indicate the 5th and 95th percentiles.

7.5 Conclusion

The application of a multiple isotope approach (δ^{15}N- and δ^{18}O-NO_3^-, and δ^{11}B) indicated seasonal predominant NO_3^- sources. Manure was the dominant source for classes A, AGC and AH both in winter and summer. Sewage was the dominant source for class H. The N isotopes are used for separating manure and sewage from mineral fertilizers and boron isotopes are a useful discriminating source indicator for separating manure from sewage based on unique isotope values. Furthermore, the SIAR output gave insight into temporal and spatial variations of both the dominant NO_3^- source and other important potential NO_3^- source contributions. Since source contributions were variable within and between sampling points, some sampling points in the same class behaved quite differently based on expert knowledge. Therefore, based on the SIAR fingerprint output, expert classification has been retrieved via a k-means clustering approach. Four classes, both for winter and summer, were retained: class A (agriculture), class AGC (agriculture with groundwater compensation), class G (greenhouses in an agricultural area), and class H (households). The expert class AH was removed as sampling points in this class were either clustered in class A or class AGW. However, the removal of class AH from the expert classification requires precaution, as the multiple isotope approach cannot exactly separate class AH from classes A and AGW. Finally, comparison of decision tree models using two year physico-chemical data and physico-chemical data with stable isotope data of NO_3^- compared to classification based on expert knowledge and k-means clustering indicates that a stable isotope fingerprint of NO_3^- helps to develop a NO_3^- source classification in absence of expert knowledge. Once the NO_3^- classes are defined either by experts or as the result of a clustering process, it is crucial to focus on improving the performance of decision tree model using only physico-chemical data as this will allow to correctly classify ca. 800 sampling points from the MAP monitoring network into one of the 4 classes: A, AGC, G and H. In future perspectives, decision tree models can be improved by using the long-term data set from VMM database, applying the best option for missing data replacement, and conducting different time series techniques for extracting feature information from the data set.

Chapter 8:

Conclusions and future research perspectives

Anthropogenic activities have severely accelerated the N cycle and increased N load in ecosystems, which gave rise to e.g. a global environmental concern for NO_3^- contamination in surface- and ground water. The major anthropogenic processes causing NO_3^- release include: overuse of N-based organic and inorganic fertilizers (Smil, 1999), elevated atmospheric N deposition (Benkovitz et al., 1996), and discharge from septic tanks and sewage systems.
The objective of this thesis was to identify and to compare the classification of NO_3^- sources in surface water in Flanders by expert knowledge and by isotopic fingerprinting and physico-chemical properties. The study sites were composed of 47 sampling points selected from the MAP monitoring network and were classified into five NO_3^- source classes based on land uses by experts: agriculture (class A), agriculture with groundwater compensation (class AGC), a combination of agriculture with horticulture (class AH), greenhouses in an agricultural area (class G) and households (class H).

MAJOR RESEARCH FINDINGS

Selection and improvement of an analytical technique for determination of $\delta^{15}N$- and $\delta^{18}O$-NO_3^-

The $AgNO_3$ method and the bacterial denitrification method are two analytical techniques for determination of $\delta^{15}N$- and $\delta^{18}O$-NO_3^- in aqueous samples. Comparison of results of 42 real surface water samples demonstrated that the $AgNO_3$ method and the bacterial denitrification method are statistically comparable, despite the fact that occasionaly large differences might occur for some specific samples.The large difference could result from a complex water sample matrix, with the presence of antibiotics, heavy metals, pesticides, NO_2^- (converted together with NO_3^- by the bacterial denitrification method) etc., affecting the bacterial denitrification method. Since the $AgNO_3$ method is relatively labor- and cost-intensive and not suitable for seawater samples, KCl extracts and freshwater samples with low NO_3^- concentration, we used the bacterial denitrification method to measure water samples of the 30 sampling points. The overall uncertainties, which are composed of the machine error and the method error,allows using this method for source identification of NO_3^-. However, a slight offset compared to true values is observed. The "bacterial denitrification method" has the following advantages: (1) sample preparation procedure is less labor-intensive (2-3 days for sample preparation) and less expensive; (2) the technique requires a smaller sample size (three orders of magnitude smaller than the "$AgNO_3$") and allows for the analysis of low NO_3^-

concentration samples; and (3) the method achieves a high level of sensitivity. Although, the overall uncertainty of the bacterial denitrification method ranges from from 0.3 to 2.2‰ for $\delta^{15}N\text{-}NO_3^-$ and from 0.8 to 2.5‰ for $\delta^{18}O\text{-}NO_3^-$, the wider isotope ranges of the 5 potential nitrate sources allows identifying nitrate contamination sources in surface water. Thus, the bacterial denitrification method is a fully operational method for the N and O isotopic analysis of nitrate.

In addition to method error and machine error, sampling error is also important, and cannot absolutely be excluded. However, if the sampling uncertainty exists in our samples, it would be the same for all water samples collected by VMM, as experts from VMM performed sampling based on their usual sampling strategy for sample collection, transport and storage.

Multiple isotope approach ($\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ and $\delta^{11}B$) for NO_3^- source identification

The NO_3^- sources, fertilizer, manure, human sewage, soil N and atmospheric N deposition are main potential sources for surface water in Flanders. The application of a multiple isotope approach ($\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$, and $\delta^{11}B$) in this study allowed to identify the predominant NO_3^- sources. Manure was the dominant source for classes A, AGC and AH both in winter and summer. Sewage was the dominant source for class H. Although $\delta^{11}B$ data indicated mineral fertilizer could be the dominant source for class G, we cannot exclude considerable contribution from NO_3^- in precipitation in summer season and manure in winter season, which was supported by the classical $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ bi-plot approach. Although we could obtain meaningful information regarding seasonal dominant sources, this is only qualitative information and there is no further quantitative message of the other potential NO_3^- source inputs, which are likely to occur in these complex conditions.

Bayesian isotope mixing model (SIAR) for estimation of NO_3^- source apportionment

Our study showed that SIAR outputs give insight into temporal and spatial variations of NO_3^- sources, which could not be achieved by the classical dual isotope bi-plot approach. SIAR offers a number of advantages as it can incorporate sources of uncertainty, isotope fractionation and multiple NO_3^- sources. However, its resolution is largely determined by:

(1) the uncertainty on the isotopic composition of the different NO_3^- sources. $\delta^{15}N$- and $\delta^{18}O\text{-}NO_3^-$ values of the five potential NO_3^- sources have relatively wide ranges and show overlap for NH_4^+ in fertilizer and rain, soil N and manure and sewage. Thus,

small variations in isotope values of NO_3^- in surface water samples might result in large changes in source apportionment as estimated by SIAR.

(2) the temporal variability of the isotopic composition of NO_3^- in the mixture. The NO_3^- source contribution may vary over time for the 30 sampling points, which caused the isotopic composition of NO_3^- in water samples to vary as well. Furthermore, the mean $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of the five potential NO_3^- sources also have large standard deviations (Xue et al., 2009). As a result, the SIAR outputs yields a wide range for individual source contributions.

Although the SIAR output provided large proportional contribution ranges for the 30 sampling points, clear seasonal patterns and source contribution patterns could be observed.

Can expert classification be retrieved via a stable isotope fingerprint?

The SIAR fingerprint indicated that source contributions were variable within and between sampling points. Thus, sampling points in the same class also behaved differently. We aimed at retrieving expert classification of the 30 sampling points based on the SIAR fingerprint using a k-mean clustering approach. The Rand index demonstrates that the k-means clustering approach based on SIAR output retrieved 70% of expert classification. Thus expert knowledge could be retrived satisfactorly. This also indicates that the "coupled SIAR – k-means clustering approach" is a promising tool for classifying nitrate contamination when relevant expert knowledge is absent. The latter would avoid long term monitoring for NO_3^- concentration (the routine method) and reduce time and labor cost during large scale sampling campaings to rertrieve NO_3^- classes in non-characterized river basins. Considering relatively large temporal and spatial variations of nitrate isotope values in this study, this is a major factor resulting in a wide range of proportional source contributions estimated by SIAR. The wide isotope ranges of source values collected from literature are another major constrained. However, these two factors can influence any mixing model. Although situation was very complex in our study area: noise data, spatio-temporal variation, overlaps of isotope ranges between sources and complex mixing of sources, the SIAR model takes into account the incorporation of source uncertainties to provide reasonable proportional source contribution ranges.

Once the NO_3^- classes are defined either by experts or as a result of a clustering process or as a combination of both methods, it is crucial to focus on improving the performance of decision tree model using only physico-chemical data as this will allow to correctly classify other

sampling locations via a classical physic-chemical monitoring/data (e.g. the ca. 750 remaining sampling points from the MAP monitoring network).

FUTURE PERSPECTIVES

Improve decision tree model performance

- The decision tree model built in this study is a univariate decision tree. Which means that at each node, a split of the data is made based on the value of a single parameter. There exist also multivariate decision trees, which split data based on the value of a linear combination of several parameters (Li et al. 2003). Furthermore, information gain is used as a splitting rule in this study, which tends to find splits where some or as many classes as possible are divided perfectly or nearly perfectly. Extra splitting rules can also be applied e.g.: (a) Gini splitting rule which searches for the largest class in the data set and strives to isolate it from the other classes and (b) twoing splitting rule which separates the classes into two groups, attempting to find groups that together added up to 50% of the data. Comparison of univariate- and multivariate decision trees with different splitting rules needs to be done in future.
- Missing data occurred in the dataset, as all of the physico-chemical parameters were not always measured at the same time. The "average replacement technique" was used to solve the missing data problem. However, replacing the missing values with a single value changes the distribution of that variable by decreasing its variance that is likely present (Pigott, 2001). Hence the data values have possibility to deviate away from the true data values. Multiple imputation (MI) has become a much more popular imputation technique, which involves replacing the missing value with 2 or more imputed values (Little and Rubin, 2002). In multiple imputation, missing values for any parameter are predicted using existing values from all other parameters in the dataset. The predicted values, called "imputes", substitute the missing values, resulting in a full dataset. This process is performed multiple times ($m > 1$), producing multiple imputed data sets (hence called "multiple imputation"). Finally missing values are substituted by the average of the simulated values over the m times. Since the physico-chemical parameters in the data set varies in function of time, it is of interest to find missing data replacement techniques which can deal with missing data via time series analyses. There is one technique called "singular spectrum analyses" (SSA), which can simultaneously fill in missing data based on extraction of additive components of time series such as trends and periodic components. This technique decomposes time series

with missing values followed by extracting trends and periodic components and then reconstruct the time series based on the extracted components.

- Another crucial step that can improve decision tree model performance is feature extraction from a data set. When a data set is composed of a large number of parameters and some of the parameters are correlated with one another, it is useful to use a mathematical technique, principal component analysis (PCA). The PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables.This technique allows to discover or to reduce the dimensionality of the data set and to identify new meaningful parameters. In addition, from the other hand, it is valuable to create new attributes that can capture the important information in a data set much more efficiently than the original attributes. A useful mathematic method is the Fourier transform, which can transform data varying in space or time into a different domain called the frequency domain. The Fourier transform is used to transform the data to the frequency domain and the inverse Fourier transform is used to reconstruct the data set which can capture the main properties contained inside for building a decision tree model.

SIAR model improvement

SIAR is a useful approach to estimate temporal and spatial variations of different NO_3^- source contributions. However, its resolution is largely determined by the temporal variability of the isotopic composition of NO_3^- in the mixture and the uncertainty on the isotopic composition of the different NO_3^- sources. Thus, to improve the SIAR model performance on NO_3^- source apportionments, it is crucial to emphasize on the following aspects in future:

- In situ estimation of fractionation factor for denitrification. Important processes in the N cycle can alter original $\delta^{15}N$- and $\delta^{18}O$-NO_3^- source values. Take an example, the ammonification process (organic-N → NH_4^+) results in a small fractionation (±1‰) between (soil) organic matter and (soil) NH_4^+ (Kendall, 1998). In contrast, the conversion of NH_4^+ to NO_2^- and NO_3^- is accompanied by marked N isotope fractionation effects (enrichment factors are between -12‰ to -29‰) (DiSpirito and Hooper, 1986; Kendall, 1998; Kool et al., 2007). Nitrogen isotope enrichment factors for denitrification are within -40‰ to -5‰ (Hübner, 1986; Smith et al., 1991; Sebilo et

al., 2003), and oxygen isotope enrichment factors are between -18‰ and -8‰ (Böttcher et al., 1990; Mengis et al., 1999; Fukada et al., 2003; Lehmann et al., 2003). Thus, a fractionation factor should be considered for the entire transformation process from release of the corresponding NO_3^- source until reaching the mixture. In situ measurements are then applied to assess denitrification in surface water.

- Properly narrow potential NO_3^- source ranges. The mean $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values of the five potential NO_3^- sources showed large standard deviations (Xue et al., 2009). Consequently, the SIAR outputs yielded a wide range for individual source contributions (see Chapters 6 and 7). The SIAR outputs might be improved by using original source material which could provide narrower ranges of the isotopic composition of NO_3^- sources. However, this could also constrain NO_3^- source apportionment since source sampling is time consuming and may also show temporal variation in its isotopic compositions.

Classify remaining sampling points from the MAP monintoring network

Our study demonstrates that the k-means clustering approach based on SIAR output retrieved 70% of expert classification. This indicates that the "coupled SIAR – k-means clustering approach" is a promising tool for classifying nitrate contamination when relevant expert knowledge is absent. This method avoid long term monitoring for NO_3^- concentration (the routine method) and reduce time and labor cost during large scale sampling campaings to rertrieve NO_3^- classes in non-characterized regions.

Once the NO_3^- classes are defined either by experts or as a result of a clustering process or as a combination of both methods, it is crucial to focus on improving the performance of decision tree model using only physico-chemical data as this will allow to correctly classify other sampling locations via a classical physic-chemical monitoring/data (e.g. the ca. 750 remaining sampling points from the MAP monitoring network).

Summary

In natural systems, the main N source are bacteria that fix N_2 from air. In non-natural systems, on the other hand, anthropogenic activities have severely changed the N cycle and increased the N load of ecosystems, often leading to high NO_3^- concentrations in surface- and ground water. The objective of this thesis was to identify and to compare the classification of NO_3^- sources in surface water in Flanders by expert knowledge via isotopic fingerprinting and physico-chemical properties. To achieve the objective, an integrated approach has been set-up by (1) selecting representative sampling points based on a decision tree model performance, (2) selecting an analytical technique for determination of $\delta^{15}N$- and $\delta^{18}O$-NO_3^-, and (3) retrieving the expert classification of 30 sampling points using outputs of a Bayesian isotope mixing model in a k-means clustering approach and performance of the decision tree model as built up in step 1.

The study sites were located in Flanders, the northern part of Belgium. Forty-seven sampling points were selected from the MAP (Manure Action Plan) monitoring network and classified by experts into 5 different NO_3^- source classes: 7 sampling points were classified into the agriculture (class A) class; 15 sampling points into the agriculture class with groundwater compensation (class AGC); 6 sampling points into the class combining agriculture with horticulture (class AH); 11 sampling points into the class with greenhouses in an agricultural area (class G) and 8 sampling points into the household (class H) class.

Chapter 2 summarizes typical $\delta^{15}N$- and $\delta^{18}O$-NO_3^- ranges of known NO_3^- sources, interprets constraints and future outlooks to quantify NO_3^- source contributions, and describes two frequently used analytical techniques ("$AgNO_3$ method" and "bacterial denitrification method") for $\delta^{15}N$- and $\delta^{18}O$-NO_3^- determination. This chapter provides a comprehensive review on N and O isotope ratios for NO_3^- source identification.

The second section (Chapter 3) is focused on the use of a decision tree model, based on long-term physico-chemical data of the 47 selected sampling points which were classified into the five NO_3^- source classes. The decision tree model learned that the classification of all sampling points was for 82% similar to the expert classification. The classification of the sampling points of classes A, AGC, G, and H were for more than 80% the same as the expert classification, while for those of class AH only 58% was correctly classified. The sampling points in class AH may have a higher probability to be labeled into another class. Based on

the decision tree model performance, 30 representative sampling points (six sampling points per class) were selected for further isotope monitoring and retrieve of the expert classification. The third section (Chapters 4 and 5) deals with the selection of a suitable method, either the silver nitrate ($AgNO_3$) or the bacterial denitrification method, to measure $\delta^{15}N$- and $\delta^{18}O$-NO_3^- of 30 sampling points, monitored from October 2007 to September 2009. A thorough method comparison (Chapter 4) was carried out using a relatively large number of surface water samples (42 water samples) with a wide range of $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values. A positive linear relation with a high correlation coefficient ($r \geq 0.88$) between both methods was found for $\delta^{15}N$- as well as for $\delta^{18}O$-NO_3^-. The comparability of both methods was assessed by the Bland-Altman technique using 95% limits of agreement. Results showed that the $AgNO_3$ and the bacterial denitrification methods are statistically interchangeable, despite the fact that the bacterial denitrification method tends to underestimate $\delta^{15}N$-NO_3^- and overestimate $\delta^{18}O$-NO_3^- values, compared to the $AgNO_3$ method. Since the $AgNO_3$ and the bacterial denitrification methods are statistically interchangeable and the low cost for the latter method, the bacterial denitrification method was considered as a suitable method. In Chapter 5, we also demonstrated the correct way to compute uncertainties on corrected $\delta^{15}N$- and $\delta^{18}O$-NO_3^- values via the bacterial denitrification method. Because errors can potentially be produced during preparation of the samples, changing analytical conditions and running equipment, it is necessary to consider an uncertainty on the final result after correction.

The last section (Chapters 6 and 7) targets on retrieving the expert classification of the 30 sampling points using outputs of a Bayesian isotope mixing model (SIAR) in a k-means clustering approach. In Chapter 7, we first applied a multiple isotope approach ($\delta^{15}N$- and $\delta^{18}O$-NO_3^- and $\delta^{11}B$) for identification of the seasonally dominant NO_3^- sources for the sampling points. Manure was the dominant source for classes A, AGC and AH both in winter and summer. Sewage was the dominant source for class H. Boron isotopes are useful discriminating source indicators for separation of manure from sewage based on unique isotope values. Furthermore, SIAR outputs gave useful insight into temporal and spatial variations of both the dominant NO_3^- source and the other important potential NO_3^- source contributions. The sampling points in classes A, AGC, AH and G were clustered using the k-means clustering approach. Since $\delta^{11}B$ showed that sewage was the dominant source for class H, this class was not considered by k-means clustering. For the remaining classes we suggested that 3 clusters were suitable for winter and summer, as it showed the same silhouette value (0.6) and Rand index (0.7) as for 4 clusters and kept the majority of the sampling points within the same expert class. As a result, we concluded to classify the

sampling points into four classes both in winter and summer: classes A, class AGC, class G, and class H as compared to the expert classification. Based on isotope characterization, the expert class AH was not retained.

In conclusion, decision tree models, built on long-term physico-chemical data and the k-means clustering approach using the two-year data for isotope and physico-chemical properties, showed similar CCI's. This indicates that stable isotope singnatures of NO_3^-, when run through an appropriate mixing model and clustering technique can help to develop a NO_3^- classification in the absence of expert knowledge for a certain basin. Once NO_3^- classes are defined in this way more sampling points could be classified via a decision tree approach using physicohemiocal data of the additional sampling points

Samenvatting

In natuurlijke systemen zijn bacteriën de belangrijkste bron van N via biologische N_2 fixatie. In niet-natuurlijke systemen daarentegen, hebben menselijke activiteiten de N-cyclus sterk gewijzigd door een verhoogde N toevoer in ecosystemen, wat vaak leidt tot hoge NO_3^- concentraties in oppervlakte- en grondwater. Het doel van dit proefschrift is de classificatie van NO_3^- bronnen in het oppervlaktewater in Vlaanderen door experten te toetsen via een classificatie bekomen via isotopische fingerprinting en fysico-chemische eigenschappen. Om deze doelstelling te bereiken werd een geïntegreerde aanpak opgezet via (1) de selectie van representatieve bemonsteringspunten op basis van het resultaat van een 'decision tree' model, (2) het selecteren van de juiste methode voor de bepaling van $\delta^{15}N$- en $\delta^{18}O$-NO_3^-, en (3) het verifiëren van de expertenclassificatie van 30 meetpunten via het gebruik van de resultaten van een Bayesiaans 'isotope mixing' model d.m.v. een 'k-means clustering' aanpak en een 'decision tree' model zoals ontwikkeld in stap 1.

De studie sites bevinden zich in Vlaanderen, het noordelijke deel van België. Zevenenveertig meetpunten werden geselecteerd uit het MAP (Mest Actie Plan) meetnet en ingedeeld in 5 verschillende NO_3^--klassen door experten: 7 meetpunten werden ingedeeld in de landbouw klasse (klasse A), 15 meetpunten in de landbouw klasse met het grondwatercompensatie (AGC-klasse), 6 meetpunten in de klasse die landbouw en tuinbouw combineert (klasse AH), 11 meetpunten in de klasse met serres in agrarisch gebied (klasse G) en 8 meetpunten in de klasse huishoudens (klasse H).

Hoofdstuk 2 toont typische $\delta^{15}N$- en $\delta^{18}O$-NO_3^- waarden voor gekende NO_3^- bronnen, interpreteert beperkingen en de toekomstige mogelijkheden om NO_3^- bronnen te kwantificeren, en beschrijft twee veel gebruikte analytische methoden ("$AgNO_3^-$ methode" en "bacteriële denitrificatie methode") voor $\delta^{15}N$- en $\delta^{18}O$-NO_3^- bepaling. Dit hoofdstuk geeft een uitvoerig overzicht van N en O isotopen ratios voor NO_3^- bronidentificatie.

Het tweede deel (Hoofdstuk 3) is gericht op het gebruik van een 'decision tree' model gebaseerd op lange termijn fysico-chemische gegevens van de 47 geselecteerde meetpunten. Het 'decision tree' model toonde dat de classificatie, van alle bemonsteringspunten samen, voor 82% overeenkwam met de expertclassificatie. De classificatie van de bemonsteringspunten van de klassen A, AGC, G en H kwam voor meer dan 80% overeen met de expertclassificatie. Klasse AH had het laagste percentage (58%) van overeenkomst. De meetpunten in klasse AH hebben een hogere kans om te belanden in een andere klasse. Dertig

representatieve bemonsteringspunten (zes meetpunten per klasse) werden geselecteerd op basis van de prestaties van het 'decision tree' model voor verdere isotopenmonitoring en verificatie van de expertclassificatie.

Het derde deel (Hoofdstukken 4 en 5) had tot doel een geschikte methode te selecteren, ofwel de zilvernitraat ($AgNO_3$) methode of de bacteriële denitrificatie methode, voor de bepaling van $\delta^{15}N$- en $\delta^{18}O$-NO_3^- voor de 30 meetpunten tijdens een controle periode van oktober 2007 tot september 2009. Een grondige vergelijking (Hoofdstuk 4) van beide methoden werd uitgevoerd met behulp van een relatief groot aantal monsters van oppervlaktewater (42 watermonsters), met een brede range van $\delta^{15}N$- en $\delta^{18}O$-NO_3^- waarden. Een positieve lineaire relatie met een hoge correlatiecoëfficiënt ($r \geq 0.88$) tussen beide methoden werd gevonden voor zowel $\delta^{15}N$- als $\delta^{18}O$-NO_3^-. De vergelijkbaarheid van beide methoden werd nagegaan via de Bland-Altman techniek met een 95% overeenkomstsgrens. De resultaten tonen dat de $AgNO_3$ en de bacteriële denitrificatie methoden statistisch uitwisselbaar zijn, ondanks het feit dat de bacteriële denitrificatie methode de neiging heeft om de $\delta^{15}N$-NO_3^- waarden te overschatten en de $\delta^{18}O$-NO_3^- waarden te onderschatten in vergelijking met de $AgNO_3$ methode. Daar beide methoden statistisch uitwisselbaar zijn, maar de bacteriële methode de goedkoopste en eenvoudiger is, werd deze laatste als meest geschikte beschouwd. In Hoofdstuk 5, hebben wij ook de correcte manier aangetoond om onzekerheden te berekenen op gecorrigeerde $\delta^{15}N$- en $\delta^{18}O$-NO_3^- waarden via de bacteriële denitrificatie methode. Aangezien fouten mogelijks worden geproduceerd tijdens de monstervoorbereiding, bij wijzigende analytische condities en apparatuur, is het noodzakelijk om de onzekerheid omtrent het uiteindelijke resultaat na correctie in ogenschouw te nemen.

Het laatste deel (Hoofdstukken 6 en 7) is gericht op de verificatie van de expertenclassificaties van de 30 meetpunten via het gebruik van de resultaten van een Bayesiaans 'isotope mixing' model (SIAR) d.m.v. een 'k-means clustering' aanpak. In Hoofdstuk 7 hebben we een meervoudige isotopenaanpak ($\delta^{15}N$- en $\delta^{18}O$- NO_3^- en $\delta^{11}B$) toegepast voor de identificatie van de dominante NO_3^- bron voor alle 30 meetpunten per seizoen. Mest was de belangrijkste bron voor de klassen A, AGC en AH zowel in de winter als de zomer. Afvalwater was de belangrijkste bron voor klasse H. Boorisotopen zijn nuttige discriminerende indicatoren voor het onderscheid tussen mest en afvalwater op basis van unieke isotopenwaarden. Bovendien, gaven de SIAR resultaten een bruikbaar inzicht in de ruimtelijke en temporele variaties van zowel de dominante NO_3^- bronnen als andere belangrijke bijdragen van potentiële NO_3^- bronnen. De meetpunten in de klassen A, AGC, AH en G werden geclusterd met behulp van een 'k-means clustering' aanpak. Daar $\delta^{11}B$ data aantoonde dat afvalwater de dominante bron

was voor klasse H, werd deze klasse niet behandeld via de ‘k-means clustering’. We suggereren dat 3 clusters geschikt waren voor de winter en de zomer, daar ze dezelfde silhouet waarde (0,6) en Rand-index (0,7) hadden als voor 4 clusters. De meerderheid van de meetpunten bleef zo in dezelfde expertenklasse. Daarom werd besloten de bemonsteringspunten in vier klassen onder te verdelen, zowel voor de winter als voor de zomer: de klassen A, klasse AGC, klasse G, en klasse H, zoals ook bij de expertenclassificatie. Op basis van isotopenkarakterisatie werd de expertenklasse AH behouden.

Tot slot kan worden gesteld worden dat ‘decision tree’ modellen, gebaseerd op, enerzijds, lange termijn fysicochemische data en, anderzijds, op data van twee jaar voor isotopen en fysico-chemische eigenschappen, vergelijkbare classificatie percentages aangeven. Dit toont aan dat isotopen kunnen helpen bij de classificatie van NO_3^- bronnen in afwezigheid van expertenkennis voor een bepaald revier bekken. Een de NO_3^- klassen op deze manier worden geïdentificeerd zouden bijkomende bemonsteringspunten kunnen worden geklasseerd via een ‘decision tree’ analyse van de physico-chemische data van de addtionele sampling points

References

Amberger, A. and Schmidt, H.L. (1987) Natural isotope contents of nitrate as indicators for its origin. Geochim. Cosmochim. Acta 51, 2699-2705.

Anderson, I.C. and Levine, J.S. (1986) Relative rates of nitric oxide and nitrous oxide production by nitrifiers, denitrifiers, and nitrate respirers. Appl. Environ. Microbiol. 51, 938-945.

Andersson, K.K., Philson, S.B. and Hooper, A.B. (1982) ^{18}O isotope shift in ^{15}N NMR analysis of biological N-oxidations: H_2O-NO_2^- exchange in the ammonia-oxidizing bacterium Nitrosomonas. Proc. Natl. Acad. Sci. 79, 5871-5875.

Aravena, R. and Robertson, W.D. (1998) Use of multiple isotope tracers to evaluate denitrification in ground water: study of nitrate from a large-flux septic system plume. Ground Water 36, 975-982.

Aravena, R., Evans, M.L. and Cherry, J.A. (1993) Stable isotopes of oxygen and nitrogen in source identification of nitrate from septic systems. Ground Water 31, 180-186.

Barnes, R.T., Raymond, P.A. and Casciotti, K.L. (2008) Dual isotope analysis indicate efficient processing of atmospheric nitrate by forested watersheds in the northeastern U.S.. Biogeochemistry 90, 15-27.

Basset, R.L., Buszka, P.M., Davidson, G.R. and Chong-Diaz, D. (1995) Identification of groundwater solute sources using boron isotopic composition. Environ. Sci. Technol. 29, 2915–2922.

Bassett, R.L. (1990) A critical evaluation of the available measurements for the stable isotopes of boron. Appl. Geochem. 5, 541–554.

Batchelor, B. and Lawrence, A.W. (1978) A kinetic model for autotrophic denitrification using elemental sulfur. Water Res. 12, 1075-1084

Bateman, A.S. and Kelly, S.D. (2007) Fertilizer nitrogen isotope signatures. Isot. Environ. Health Stud. 43, 237-247.

Benkovitz, C.M., Scholtz, M.T., Pacyna, J., Tarrason, L., Dignon, J., Voldner, E.C., Spiro, P.A., Logan, J.A. and Graedel, T.E. (1996) Global gridded inventories of anthropogenic emissions of sulfur and nitrogen. Geophys. Res. 101, 29239-29253.

Black, A.S. and Waring, S.A. (1977) The natural abundance of ^{15}N in the soil-water system of a small catchment area. Aust. J. Soil Res. 15, 51-57.

Bland, J.M. and Altman, D.G. (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet i, 307–310.

Böhlke, J.K., Ericksen, G.E. and Revesz, K. (1997) Stable isotope evidence for an atmospheric origin of desert nitrate deposits in northern Chile and southern California, U.S.A. Chem. Geol. 136, 135- 152.

Böttcher, J., Strebel, O., Voerkelius, S. and Schmidt, H.L. (1990) Using isotope fractionation of nitrate-nitrogen and nitrate-oxygen for evaluation of microbial denitrification in a sandy aquifer. J. Hydrol. 114, 413-424.

Brand, W.A.; Coplen, T.B.; Aerts-Bijma, A.T.; Böhlke, J.K.; Gehre, M.; Geilmann, H.; Gröning, M.; Jansen, H.G.; Meijer, H.A.J.; Mroczkowski, S.J.; Qi, H.; Soergel, K.; Stuart-Williams, H.; Weise, S.M. and Werner, R.A. (2009) Comprehensive inter-laboratory calibration of reference materials for $\delta^{18}O$ versus VSMOW using various on-line high-temperature conversion techniques. Rapid Commun. Mass Spectrom. 23, 999-1019.

Brandes, J. A., Devol, A. H., Yoshinari, T., Jayakumar, D. A. and Naqvi, S.W.A. (1998) Isotopic composition of nitrate in the central Arabian Sea and eastern tripical North Pacific: a tracer for mixing and nitrogen cycles. Limnol. Oceanogr. 43, 1680-1689.

Burg, A. and Heaton, T.H.E. (1998) The relationship between the nitrate concentration and hydrology of a small chalk spring; Israel. J. Hydrol. 204, 68-82.

Cao, Y.C., Sun, G.Q., Xing, G.X. and Xu, H. (1991) Natural abundance of ^{15}N in main N-containing chemical fertilizers of China. Pedosphere 1, 377-382.

Casciotti, K.L., Sigman, D.M., Galanter Hastings, M., Böhlke, J.K. and Hilkert, A. (2002) Measurement of the oxygen isotopic composition of nitrate in seawater and freshwater using the denitrifier method. Anal. Chem. 74, 4905-4912.

Cazaux, G., Carels, K. and Van Gijseghem, D. (2007) Prospects and challenges for agricultural diversification in a peri-urban region (Flanders-Belgium). August 2007.

Chang, C.C.Y., Langston, J., Riggs, M., Campbell, D.H., Silva, S.R. and Kendall, C. (1999) A method for nitrate collection for ^{15}N and ^{18}O analysis from waters with low nitrate concentrations. Can. J. Fish. Aquat. Sci. 56, 1856-1864.

Chetelat, B. and Gaillardet, J. (2005) Boron isotopes in the Seine River, France: a probe of anthropogenic contamination. Environ. Sci. Technol. 39, 2486–2493.

Choi, W.J., Han, G.H., Lee, S.M., Lee, G.T., Yoon, K.S., Choi, S.M. and Ro, H. M. (2007) Impact of land-use types on nitrate concentration and $\delta^{15}N$ in unconfined ground water in rural areas of Korea. Agric. Ecosyst. Environ. 120, 259 -268.

Choi, W.J., Han, G.H., Ro, H.M., Yoo, S.H. and Lee, S.M. (2002a) Evaluation of nitrate contamination sources of unconfined groundwater in the North Han River basin of Korea using nitrogen isotope ratios. Geosciences 6, 47-55.

Choi, W.J., Lee, S.M. and Ro, H.M. (2003a) Evaluation of contamination sources of ground water NO_3^- using nitrogen isotope data: A review. Geosciences 7, 81-87.

Choi, W.J., Lee, S.M., Ro, H.M., Kim, K.C. and Yoo, S.H. (2002b) Natural ^{15}N abundances of maize and soil amended with urea and composted pig manure. Plant and Soil 245, 223-232.

Choi, W.J., Ro, H.M. and Lee, S.M. (2003b) Natural ^{15}N abundances of inorganic nitrogen in soil treated with fertilizer and compost under changing soil moisture regimes. Soil Biol. Biochem. 35, 1289-1298.

Cook, G.A. and Lauer, C. M. (1968) "Oxygen", in Clifford A. Hampel: The Encyclopedia of the Chemical Elements. New York: Reinhold Book Corporation, 499-512. LCCN 68-29938.

Curt, M.D., Aguado, P., Sánchez, G., Bigeriego, M. and Fernández, J. (2004) Nitrogen isotope ratios of synthetic and organic sources of nitrate water contamination in Spain. Water, Air and Soil Pollution 151, 135-142.

Delwiche, C., and Steyn, P. (1970) Nitrogen isotope fractionation in soils and microbial reactions. Environ. Sci. Technol. 4, 929-935.

Deutsch, B., Kahle, P. and Voss, M. (2006b) Assessing the source of nitrate pollution in water using stable N and O isotopes. Agron. Sustain. Dev. 26, 263-267.

Deutsch, B., Liskow, I., Kahle, P. and Voss, M. (2005) Variation in the $\delta^{15}N$ and $\delta^{18}O$ values of nitrate in drainage water of two fertilized fields in Mecklenburg-Vorpommern (Germany). Aquat. Sci. 67, 156-165.

Deutsch, B., Mewes, M., Liskow, I. and Voss, M. (2006a) Quantification of diffuse nitrate inputs into a small river system using stable isotopes of oxygen and nitrogen in nitrate. Org. Geochem. 37, 1333-1342.

DiSpirito, A.A. and Hooper, A.B. (1986) Oxygen exchange between nitrate molecules during nitrite oxidation by *Nitrobacter*. Biol. Chem. 261, 10534-10537.

Du, W.L. & Zhan, Z.J. (2002) Conferences in research and practice in information technology. Vol. 14.

Durka, W., Schulze, E.D., Gebauer, G. and Voerkelius, S. (1994) Effects of forest decline on uptake and leaching of deposited nitrate determined from ^{15}N and ^{18}O measurements. Nature 372**,** 765-767.

EC, 2002. Implementation of Council Directive 91/676/EEC concerning the protection of waters against pollution caused by nitrates from agricultural sources.

Eisenhut, S. and Heumann, K.G. (1997) Identification of ground water contaminations by landfills using precise boron isotope ratio measurements with negative thermal ionization mass spectrometry. Fresen. J. Anal. Chem. 359, 375-377.

Ellison, A.M. (2004) Bayesian inference in ecology. Ecol. Lett., 7, 509-520.

Eppinger, R., Vandevelde, D, Dobbelaere, A. and Maeckelberghe, H. (2005) Monitoring effectiveness of the EU Nitrates Directive Action Programmes: Approach by the Flemish region (Belgium). RIVM report 500003007/2005: 57-83.

European Environment Agency. Source apportionment of nitrogen and phosphorus inputs into the aquatic environment. EEA Report No 7/2005, ISSN 1725-9177.

Evans, M., Hastings, N. & Peacock, B. (2000) Statistical Distributions, 3rd edn. John Wiley and Sons, New York.

Feast, N.A., Hiscock, K.M., Dennis, P.F. and Andrews, J.N. (1998) Nitrogen isotope hydrochemistry and denitrification within the Chalk aquifer system of north Norfolk, UK. J. Hydrol. 211, 233-252.

Fielding, S., Fayers, P.M., Loge, J.H., and Kaasa, S. (2006) Methods for handling missing data in palliative care. Palliative Medicine 20, 791-798.

Finlay, J.C., Sterner, R.W. and Kumar, S. (2007) Isotopic evidence for in-lake production of accumulating nitrate in Lake Superior. Ecol. Appl. 17, 2323-2332.

Flipse, W.J. and Bonner, F.T. (1985) Nitrogen-isotope ratios of nitrate in ground water under fertilized fields, Long Island, New York. Ground Water 23, 59-67.

Fogg, G.E., Rolston, D.E., Decker, D.L., Louie, D.T. and Grismer, M.E. (1998) Spatial variation in nitrogen isotope values beneath nitrate contamination sources. Ground Water 36, 418-426.

Fukada, T., Hiscock, K.M. and Dennis, P.F. (2004) A dual isotope approach to the nitrogen hydrochemistry of an urban aquifer. Appl. Geochem. 19, 709-719.

Fukada, T., Hiscock, K.M., Dennis, P.F. and Grischek, T. (2003) A dual isotope approach to identify denitrification in ground water at a river bank infiltration site. Water Res. 37, 3070-3078.

Gellenbeck, D.J. (1994) Isotopic compositions and sources of nitrate in ground water from western Salt River Valley, Arizona: U.S. Geological Survey Water-Resources Investigations Report 94–4063, 53 p.

Goldberg, S., Lesch, S.M. and Suarez, D.L. (2000) Predicting boron adsorption by soils using soil chemical parameters in the constant capacitance model. Soil Sci. Soc. Am. J. 64, 1356-1363.

Hales, H.C., Ross, D.S. and Lini, A. (2007) Isotopic signature of nitrate in two contrasting watersheds of Brush Brook, Vermont, USA. Biogeochemistry 84, 51-66.

Heaton, T.H.E. (1986) Isotopic studies of nitrogen pollution in the hydrosphere and atmosphere: a review. Chem. Geol. (Isot Geosci Section) 59, 87-102.

Hirata, K.M. (1996) Pollution of Soil and Ground Water and its Management. Tokyo: Law and Regulations Center Publishing House.

Hollocher, T.C. (1984) Source of the oxygen atoms of nitrate in the oxidation of nitrite by *Nitrobacter agilis* and evidence against a P-O-N anhydride mechanism in oxidative phosphorylation. Arch. Biochem. Biophys. 233, 721-727.

Hübner, H. (1986) Isotope effects of nitrogen in the soil and biosphere. In: Handbook of Environmental Isotope Geochemistry, eds. Fritz, P. and Fontes, J.Ch., pp. 361-425. Elsevier, Amsterdam.

IAEA. 2004. IAEA analytical quality control services reference materials catalogue 2004-2005. IAEA, Vienna, Austria.

Iqbal, M.Z., Krothe, N.C. and Spalding, R.F. (1997) Nitrogen isotope indicators of seasonal source variability to ground water. Environ. Geol. 32, 210-218.

Jaarverslag waterkwaliteit 2005, VMM, 2005.

Jackson, A.L., Inger, R., Bearhop, S. and Parnell, A. (2009) Erroneous behaviour of MixSIR, a recently published Bayesian isotope mixing model: a discussion of Moore & Semmens (2008). Ecol. Lett. 12, E1-E5.

Jardine, T.D. and Cunjak, R.A. (2005) Analytical error in stable isotope ecology. Oecologia 144, 528-533.

Johannsen, A., Dähnke, K., and Emeis, K. (2008) Isotopic composition of nitrate in five German rivers discharging into the North Sea. Organ. Geochem. 39, 1678-1689.

Jun, S.C., Bae, G.O., Lee, K.K., Chung, H.J. (2005) Identification of the source of nitrate contamination in ground water below an agricultural site, Jeungpyeong, Korea. J. Environ. Qual. 34, 804–815.

Junk, G. and Svec, H.V. (1958) The absolute abundance of the nitrogen isotopes in the atmosphere and compressed gas from various sources. Geochim. Cosmochim. Acta 14, 234-243.

Karr, J.D., Showers, W.J., Wendell Gilliam, J. and Scott Andres, A. (2001) Tracing nitrate transport and environmental impact from intensive swine farming using delta nitrogen-15. J. Environ. Qual. 30, 1163-1175.

Katz, B.G., Chelette, A.R. and Pratt, T.R. (2004) Use of chemical and isotopic tracers to assess nitrate contamination and ground-water age, Woodville Karst Plain, USA. J. Hydrol. 289, 36-61.

Kaushal, S.S., Lewis, W.M. and McCutchan, J.H. (2006) Land use change and nitrogen enrichment of a Rocky Mountain watershed. Ecol. Appl. 16, 299-312.

Kellman, L. and Hillaire-Marcel, C. (1998) Nitrate cycling in streams: using natural abundances of NO_3-δ15N to measure in-situ denitrification. Biogeochemistry 43, 273-292.

Kellman, L.M. (2005) A study of tile drain nitrate-delta N-15 values as a tool for assessing nitrate sources in an agricultural region. Nutr. Cycl. Agroecosyst. 71, 131-137.

Kellman, L.M. and Hillaire-Marcel, C. (2003) Evaluation of nitrogen isotopes as indicators of nitrate contamination sources in an agricultural watershed. Agric. Ecosyst. Environ. 95, 87-102.

Kendall, C. (1998) Tracing sources and cycling of nitrate in catchments. In: Isotope Tracers in Catchment Hydrology, eds. Kendall, C. and McDonnell, J.J., pp. 519-576. Elsevier, Amsterdam.

Kendall, C. and Aravena, R. (1999) Nitrate isotopes in groundwater systems. In: Environmental tracers in subsurface hydrology, eds. Cook, P.G. and Herczeg, A.L., pp. 261-297. Kluwer Academic Publishers.

Kendall, C. and Grim, E. (1990) Combustion tube method for measurement of nitrogen isotope ratios using calcium oxide for total removal of carbon dioxide and water. Anal. Chem. 62, 526-529.

Kendall, C., Silva, S.R., Chang, C.C.Y., Burns, D.A., Campbell, D.H. and Shanley, J.B. (1996) Use of the Delta 18-O and Delta 15-N of nitrate to determine sources of nitrate in early spring runoff in forested catchments. In Isotopes in Water Resources Management. International Atomic Energy Agency Symposium 1, 167-176.

Kish, L. (1965) Cluster sampling and subsampling. Survey Sampling. John Wiley and Sons, Inc. London.

Knowles, R. (1982) Denitrification. Microbiol. Rev. 46, 43-70.

Kohl, D. H., Shearer, G. B. and Commoner, B. (1971) Fertilizer nitrogen: Contribution to nitrate in surface water in a corn belt watershed. Science 174, 1331-1334.

Komor, S.C. (1997) Boron contents and isotopic compositions of hog manure, selected fertilizers, and water in Minnesota. J. Environ. Qual. 26, 1212-1222.

Kool, D.M., Wrage, N., Oenema, O., Dolfing, J. and Van Groenigen, J.W. (2007) Oxygen exchange between (de)nitrification intermediates and H_2O and its implication for source determination of NO_3^- and N_2O: a review. Rapid Commun. Mass Spectrom. 21, 3659-3578.

Landis, J. R.and Koch, G. G. (1977) The measurement of observer agreement for categorical data. Biometrics 33, 159–174.

Lee, K.S., Bong, Y.S., Lee, D., Kim, Y. and Kim, K. (2008) Tracing the sources of nitrate in the Han River watersheds in Korea, using $\delta^{15}N\text{-}NO_3^-$ and $\delta^{18}O\text{-}NO_3^-$ values. Sci. Total Environ. 395, 117-124.

Leenhouts, J.M., Bassett, R.L. and Maddock III, T. (1998) Utilization of intrinsic boron isotopes as co-migrating tracers for identifying potential nitrate contamination sources. Ground Water 36, 240–250.

Lefebvre, S., Clément, J.C., Pinay, G., Thenail, C., Durand, P. and Marmonier, P. (2007) ^{15}N-nitrate signature in low-order streams: effects of land cover and agricultural practices. Ecol. Appl. 17, 2333- 2346.

Lehmann, M.F., Reichert, P., Bernasconi, S.M., Barbieri, A. and McKenzie, A. (2003) Modelling nitrogen and oxygen isotope fractionation during denitrification in a lacustrine redox-transition zone. Geochim. Cosmochim. Acta 67, 2529-2542.

Lemarchand, D., Gaillardet, J., Göpel, C. and Manhès, G. (2002) An optimised procedure for boron separation and mass spectrometry analysis for river samples, Chemical Geology 182, 323-334

Lemarchand, E., Schott, J. and Gaillardet, J. (2005) Boron isotopic fractionation related to boron sorption on humic acid and the structure of surface complexes formed. Geochim. Cosmochim. Acta 69, 3519-3533.

Li, X.B., Sweigart, J.R., Teng, J.T.C., Donohue, J.M., Thombs, L.A. and Michael Wang, S. (2003) Multivariate decision trees using linear discriminants and tabu search. IEEE transactions on systems, man, and cybernetics-part A: systems and humans 33: 194-199

Li, X.D., Masuda, H., Koba, K. and Zeng, H.A. (2007) Nitrogen isotope study on nitrate-contaminated groundwater in the Sichuan Basin, China. Water, Air and Soil Pollution 178, 145-156.

Lindau, C.W., Delaune, R.D. and Alford, D.P (1997) Monitoring nitrogen pollution from sugarcane runoff using ^{15}N analysis. Water, Air and Soil Pollution 89, 389-399

Little, R.J.A. and Rubin,D.B. (2002) Statistical analysis with missing data. Hoboken, NJ: John Willey and Sons.

Macko, S.A. and Ostrom, N.E. (1994) Molecular and pollution studies using stable isotope. In: Stable Isotopes in Ecology and Environmental Science, eds. Lajtha, K. and Michner, R., pp. 45-62. Blackwell Scientific, Oxford, UK.

MacQueen, J.B. (1967) Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press 1, 281-297.

Mariotti, A., Germon, J.C., Hubert, P., Kaiser, P., Letolle, R., Tardieux, A. and Tardieux, P. (1981) Experimental determination of nitrogen kinetic isotope fractionation: some principles: illustration for the denitrification and nitrification processes. Plant and Soil 62, 423-430.

Mariotti, A., Landreau, A. and Simon, B. (1988) ^{15}N isotope biogeochemistry and natural denitrification process in ground water: application to the chalk aquifer in northern France. Geochim. Cosmochim. Acta 52, 1869-1878.

Mayer, B., Bollwerk, S.M., Mansfeldt, T., Hütter, B. and Veizer, J. (2001) The oxygen isotope composition of nitrate generated by nitrification in acid forest floors. Geochim. Cosmochim. Acta 65, 2743-2756.

Mayer, B., Boyer, E.W., Goodale, C., Jaworski, N.A., Breemen, N.V., Howarth, R.W., Seitzinger, S., Billen, G., Lajtha, K., Nadelhoffer, K., Dam, D.V., Hetling, L.J., Nosal, M. and Paustian, K. (2002) Sources of nitrate in rivers draining sixteen watersheds in the northeastern U.S.: Isotopic constraints. Biogeochemistry 57/58, 171-197.

McClelland, J.W. and Valiela, I. (1998) Linking nitrogen in estuarine producers to land-derived sources. Limnol. Oceanogr. 43, 577-585.

McGraw, K.O. and Wong, S.P. (1996) Forming inferences about some intraclass correlation coefficients. Psychological Methods 1, 30–46.

Mengis, M., Schiff, S.L., Harris, M., English, M.C., Aravena, R., Elgood, R.J. and Maclean, A. (1999) Multiple geochemical and isotopic approaches for assessing ground water NO_3^- elimination in a riparian zone. Ground Water 37, 448-457.

Mikhail, E. (1976) Observations and Least Squares. New York: IEP-A Dun-Donnelley Publisher.

Min, J.H., Yun, S.T., Kim, K., Kim, H.S., Hahn, J. and Lee, K.S. (2002) Nitrate contamination of alluvial ground waters in the Nakdong River basin, Korea. Geosci. J. 6, 35-46.

Mitchell, R.J., Scott Babcock, R., Gelinas, S., Nanus, L. and Stasney, D. (2003) Nitrate distribution and source identification in the Abbotsford-Sumas aquifer, Northwestern Washington State. J. Environ. Qual. 32, 789-800.

Moore, J.W. and Semmens, B.X. (2008) Incorporating uncertainty and prior information into stable isotope mixing models. Ecol. Lett. 11, 470-480.

Moore, K.B., Ekwurzel, B., Esser, B.K., Bryant Hudson, G. and Moran, J.E. (2006) Sources of ground water nitrate revealed using residence time and isotope methods. Appl. Geochem. 21, 1016-1029.

Nestler, A., Berglund, M, Accoe, F., Duta, S., Xue, D., Boeckx, P. and Taylor, P. (2011) Isotopes for improved management of nitrate pollution in aqueous resources: review of surface water field studies. Environ. Sci. Pollut Res. DOI 10.1007/s11356-010-0422-z

Palmer, M.R., Spivack, A.J. and Edmond, J.M. (1987) Temperature and pH controls over isotopic fractionation during adsorption of boron on marine clay. Geochim. Cosmochim. Acta 51, 2319–2323.

Panno, S.V., Hackley, K.C., Hwang, H.H. and Kelly, W.R. (2001) Determination of the sources of nitrate contamination in karst springs using isotopic and chemical indicators. Chem. Geol. 179, 113-128.

Panno, S.V., Hackley, K.C., Kelly, W.R. and Hwang, H.H. (2006) Isotopic evidence of nitrate sources and denitrification in the Mississippi River, Illinois. J. Environ. Qual. 35, 495-504.

Pardo, L.H., Kendall, C., Pett-Ridge, J. and Chang, C.C.Y. (2004) Evaluating the source of stream water nitrate using ^{15}N and ^{18}O in nitrate in two watersheds in New Hampshire, USA. Hydrol. Process. 18, 2699-2712.

Parnell, A.C., Inger, R., Bearhop, S. and Jackson, A.L. (2010) Source partioning using stable isotopes: coping with too much variation. PLoS ONE 5(3), e9672.

Phillips, D.L. and Koch, P.L. (2002) Incorporating concentration dependence in stable isotope mixing models. Oecologia 130, 114-125.

Piatek, K.B., Mitchell, M.J., Silva, S. R. and Kendall, C. (2005) Sources of nitrate in snowmelt discharge: evidence from water chemistry and stable isotopes of nitrate. Water, Air and Soil Pollution 165, 13-35.

Pidwirny, M. (2006) "The nitrogen cycle". Fundamentals of Physical Geography, 2nd Edition.

Pigott, T.D. 2001. A review of the methods for missing data. Educational Research and Evaluation 7:353-383

Piña-Ochoa, E. and Álvarez-Cobelas, M. (2006) Denitrification in aquatic environments: a cross-system analysis. Biogeochemistry 81: 111–130.

Prasad, R. and Power, J.F. (1995) Nitrification inhibitor for agriculture, health and environment. Adv. Agron. 54, 233-281.

Quinlan, J. R. (1993) C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers.

Rand, W.M. (1971) Objective criteria for the evaluation of clustering methods. J. Amer. Statistical Assoc. 66, 846–850.

Roadcap, G.S., Hackley, K.C., Hwang, H.H. and Johnson, T.M. (2001) Application of nitrogen and oxygen isotopes to identify sources of nitrate. Illinois Groundwater Consortium Conference on 9 April 2001.

Rock, L. and Ellert, B.H. (2007) Nitrogen-15 and oxygen-18 natural abundance of potassium chloride extractable soil nitrate using the denitrifier method. Soil Sci. Soc. Am. J. 71, 355-361.

Rock, L. and Mayer, B. (2002) Isotopic assessment of sources and processes affecting sulfate and nitrate in surface water and ground water of Luxembourg. Isot. Environ. Health Stud. 38, 191-206.

Rosenberg, M.S., Adams, D.C. and Gurevitch, J. (2000) Meta Win: Statistical Software for Meta-Analysis. Version 2.0. Sinauer Associates, Sunderland, Massachusetts.

Rousseeuw, P.J. (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics. 20, 53-65.

Russell, K.M., Galloway, J.N., Macko, S.A., Moody, J.L. and Scudlark, J.R. (1998) Sources of nitrogen in wet deposition to the Chesapeake Bay region. Atmos. Environ. 32, 2453-2465.

Sebilo, M., Billen, G., Grably, M. and Mariotti, A. (2003) Isotopic composition of nitrate-nitrogen as a marker of riparian and benthic denitrification at the scale of the whole Seine river system. Biogeochemistry 63, 35-51.

Seiler, R.L. (2005) Combined use of ^{15}N and ^{18}O of nitrate and ^{11}B to evaluate nitrate contamination in groundwater. Appl. Geochem. 20, 1626-1636.

Shannon, C. (1948). A mathematical theory of communication. The Bell Systems Technical Journal, 27, 379-423.

Shrout, P.E. and Fleiss, J.L. (1979) Intraclass Correlations: Uses in Assessing Rater Reliability. Psychological Bulletin 86, 420–428.

Sigman, D.M., Casciotti, K.L., Andreani, M., Barford, C., Galanter, M. and Böhlke, J.K. (2001) A bacterial method for the nitrogen isotopic analysis of nitrate in seawater and freshwater. Anal. Chem. 73, 4145-4153.

Silva, S.R., Kendall, C., Wilkison, D.H., Ziegler, A.C., Chang, C.C.Y. and Avanzino, R.J. (2000) A new method for collection of nitrate from fresh water and the analysis of nitrogen and oxygen isotope ratios. J. Hydrol. 228, 22-36.

Singleton, M.J., Esser, B.K., Moran, J.E., Hudson, G.B., McNab, W.W and Harter, T. (2007) Saturated zone denitrification: potential for natural attenuation of nitrate contamination in shallow groundwater under dairy operations. Environ. Sci. Technol. 41,759-765.

Smets, D., Carels, K., Platteau, J. and Van Gijseghem, D. (2004) Agriculture and horticulture in Flanders. Agriculture and Fisheries report.

Smil, V. (1999) Nitrogen in crop production: An account of global flows. Global Biogeochem. Cy. 13, 465-472.

Smil, V. (2000) Cycles of Life Civilization and the Biosphere. Scientific American Library, New York, 221 pp.

Smith, R.L., Howes, B.L. and Duff, J.H. (1991) Denitrification in nitrate-contaminated groundwater: occurrence in steep vertical geochemical gradients. Geochim. Cosmochim. Acta 55, 1815-1825.

Spalding,R.F., and Exner, M.E. (1993) Occurrence of nitrate in ground- water—A review. J. Environ. Qual. 22, 392-402.

Spoelstra, J., Schiff, S.L., Elgood, R.J., Semkin, R.G. and Jeffries, D.S. (2001) Tracing the sources of exported nitrate in the turkey lakes watershed using $^{15}N/^{14}N$ and $^{18}O/^{16}O$ isotopic ratios. Ecosystems 4, 536-544.

Spoelstra, J., Schiff, S.L., Hazlett, P.W., Jeffries, D.S. and Semkin, R.G. (2007) The isotopic composition of nitrate produced from nitrification in a hardwood forest floor. Geochim. Cosmochim. Acta 71, 3757-3771.

Spruill, T.B., Showers, W.J. and Howe, S.S. (2002) Application of classification-tree method to identify nitrate sources in ground water. J. Environ. Qual. 30, 1538-1549.

Tirez, K., Brusten, W., Widory, D., Petelet, E., Bregnot, A., Xue, D., Boeckx, P. and Bronders, J. (2010) Boron isotope ratio ($\delta^{11}B$) measurements in Water Framework Directive monitoring programs: comparison between double focusing sector field ICP and thermal ionization mass spectrometry. J. Anal. At. Spectrom. 25, 964-974

Torrentó, C., Cama, J., Urmeneta, J., Otero, N. and Soler, A. (2010) Denitrification of groundwater with pyrite and Thiobacillus denitrificans. Chemical Geology, 278, 80-91.

Townsend-Small, A., McCarthy, M. J., Brandes, J. A., Yang, L., Zhang, L. and Gardner, W.S. (2007) Stable isotopic composition of nitrate in Lake Taihu, China, and major inflow rivers. Hydrobiologia 581, 135-140.

Umezawa, Y., Hosono, T., Onodera, S., Siringan, F., Buapeng, S., Delinom, R., Yoshimizu, C., Tayasu, I., Nagata, T. and Taniguchi, M. (2008) Sources of nitrate and ammonium contamination in ground water under developing Asian megacities. Sci. Total Environ. 404, 361-376.

Vengosh, A., Barth, S., Heumann, K.G. and Eisenhut, S. (1999) Boron isotopic composition of freshwater lakes from Central Europe and possible contamination sources. Acta Hydrochim. Hydrobiol. 27, 416–421.

Vengosh, A., Heumann, K.G., Juraske, S. and Kasher, R. (1994) Boron isotope application for tracing sources of contamination in groundwater. Environ. Sci. Technol. 28, 1968–1974.

Verstraeten, I.M., Fetterman, G.S., Meyer, M.T., Bullen, T. and Sebree, S.K. (2005) Use of tracers and isotopes to evaluate vulnerability of water in domestic wells to septic waste. Ground Water Monit. Remediat. 25, 107–117.

Voss, M., Deutsch, B., Elmgren, R., Humborg, C., Kuuppo, P., Pastuszak, M., Rolff, C. and Schulte, U. (2006) Sources identification of nitrate by means of isotopic tracers in the Baltic Sea catchments. Biogeosciences 3, 663-676.

Wassenaar, L.I. (1995) Evaluation of the origin and fate of nitrate in the Abbotsford Aquifer using the isotopes of ^{15}N and ^{18}O in NO_3^-. Appl. Geochem. 10, 391-405.

Wells, E.R. and Krothe, N.C. (1989) Seasonal fluctuation in δ^{15} N of groundwater nitrate in a mantled karst aquifer due to macropore transport of fertilizer-derived nitrate. J. Hydrol. 112, 191-201.

Widory, D., Kloppmann, W., Chery, L., Bonnin, J., Rochdi, H. and Guinamant, J.L. (2004) Nitrate in groundwater, an isotope multi-tracer approach. J. Contam. Hydrol. 72, 165-188.

Widory, D., Petelet-Giraud, E., Négrel, P. and Ladouche, B. (2005) Tracking the sources of nitrate in groundwater using coupled nitrogen and boron isotopes, A synthesis. Environ. Sci. Technol. 39, 539 -548.

Williard, K.W.J., De Walle, D.R., Edwards, P.J. and Sharpe, W.E. (2001) ^{18}O isotopic separation of stream nitrate sources in mid-Appalachian forested watersheds. J. Hydrol. 252, 174-188.

Xue, D., Botte, J., De Baets, B., Accoe, F., Nestler, A., Taylor, P., Van Cleemput, O., Berglund, M. and Boeckx, P. (2009) Present limitations and future prospects of stable

isotope methods for nitrate source identification in surface- and groundwater. Water Res. 43, 1159–1170.

Xue, D., De Baets, B., Vermeulen, J., Botte, J., Van Cleemput, O. and Boeckx, P. (2010) Error assessment of nitrogen and oxygen isotope ratios of nitrate as determined via the bacterial denitrification method. Rapid Commun. Mass Spectrom. 24, 1979-1984.

Yingkai, X. and Lan, W. (2001). The effect of pH and temperature on the isotopic fractionation of boron between saline brine and sediments. Chem. Geol. 171, 253-261.

Žalik, K.R. (2008) An efficient k'-means clustering algorithm. Pattern Recognition Letters 29: 1385-1391.

Zhang, Y., Liu, X.J., Fangmeier, A., Goulding, K.T.W. and Zhang, F.S. (2008) Nitrogen inputs and isotopes in precipitation in the North China Plain. Atmos. Environ. 42, 1436-1448.

Curriculum Vitae

1. PERSONAL DATA

Name:	Dongmei Xue
Date of birth:	2 December 1981
Nationality:	Chinese
Address:	Sint-Denijslaan 153A, B-9000 Gent
Civil state:	Married
Mobile Phone:	+32 (0)484 781 895
E-mail address:	dongmei.xue@ugent.be, xuedongmei@gmail.com

2. EDUCATION

November 2006—Present	PhD candidate in Applied Biological Science (Thesis: Nitrate source classification in surface water via isotopic fingerprinting), Faculty of Bioscience Engineering, Ghent University, Gent, Belgium (PhD defence is expected in May 2011)
September 2005—September 2006	Master in Environmental Science and Technology (Thesis: DGT determination of total mercury and methylmercury in natural water), Faculty of Science, Vrije Universiteit Brussel (VUB), Brussels, Belgium
September, 2000—July 2004	Bachelor in Environmental Engineering (Thesis: The effects of organic acids on the release of soil silicon), Faculty of Soil and Environment, Shenyang Agricultural University, Shengyang, China

September 1998—July 2000	Fushun No10. Middle School, Liaoning Province, China

3. CAREER EXPERIENCE

July 2004—September 2005	Chemical analyst in Analysis and Measurement Centre in institute of Applied Ecology, Chinese Academy of Sciences. Responsibilities: (a) analyzing and quantifying Vitamins, Antibiotic and residues of pesticides in food using High Performance Liquid Chromatography and Mass Spectrometry (HPLC-MS); (b) measuring physical and chemical properties of natural water; and (c) participating in environmental impact assessment

4. AWARD

Chinese government award for outstanding self-financed students abroad for 2010. This award is for the students who studied abroad with excellent research work.

5. PUBLICATIONS

Publications in international journals with peer-review

Xue, D., Botte, J., De Baets, B., Accoe, F., Nestler, A., Taylor, P., Van Cleemput, O., Berglund, M. and Boeckx, P. (2009) Present limitations and future prospects of stable isotope methods for nitrate source identification in surface- and groundwater. Water Res. 43, 1159–1170.

Xue, D., De Baets, B., Botte, J., Vermeulen, J., Van Cleemput, O. and Boeckx, P. (2010) Comparison of the silver nitrate and bacterial denitrification methods for the determination of nitrogen and oxygen isotope ratios of nitrate in surface water. Rapid Commun. Mass Spectrom. 24, 833-840.

Xue, D., De Baets, B., Vermeulen, J., Botte, J., Van Cleemput, O. and Boeckx, P. (2010) Error assessment of nitrogen and oxygen isotope ratios of nitrate as determined via the bacterial denitrification method. Rapid Commun. Mass Spectrom. 24, 1979-1984.

Tirez, K., Brusten, W., Widory, D., Petelet, E., Bregnot, A., **Xue, D.,** Boeckx, P. and Bronders, J. (2010) Boron isotope ratio ($\delta^{11}B$) measurements in Water Framework Directive monitoring programs: comparison between double focusing sector field ICP and thermal ionization mass spectrometry. J. Anal. At. Spectrom. 25, 964-974

Nestler, A., Berglund, M, Accoe, F., Duta, S., **Xue, D.,** Boeckx, P. and Taylor, P. (2011) Isotopes for improved management of nitrate pollution in aqueous resources: review of surface water field studies. Environ. Sci. Pollut Res. DOI 10.1007/s11356-010-0422-z

Xue, D., De Baets, Van Cleemput, O., Hennessy, C., Berglund, M. and Boeckx, P. A Bayesian isotope mixing model to estimate proportional contributions of multiple nitrate sources in surface water. Submitted to Water Res. Under revision.

El Gaouzi, J., Sebilo, M., Boeckx, P., Ribstein, P., Plagnes, V., **Xue, D.,** Zakeossian, M. Stable isotope of oxygen and nitrogen for quantifying nitrate sources in karstic springs-Lunain watershed-basin of Paris (France). In preparation for Isotopes in environmental and Health studies.

Other publications

Xue, D., De Baets, B., Botte, J., Van Hoof, K. Boeckx, P. Combination of $\delta^{15}N$and $\delta^{18}O$ values of nitrate to identify nitrate sources in Flanders, Belgium. Benelux Association of Stable Isotope Scientists (BASIS), Arnhem, Netherlands, 3-4 April, 2008. Programme and abstracts.

Xue, D., De Baets, B., Botte, J., Van Hoof, K. and Boeckx, P. Identification and quantification of NO_3^- inputs into surface water in Flanders, Belgium. Benelux Association of Stable Isotope Scientists (BASIS), Brugge, Belgium, 2-3 April, 2009. Programme and abstracts.

Xue, D., De Baets, B., Botte, J., Van Cleemput, O. and Boeckx, P. Bayesian isotopic mixing model to quantify NO_3^- input sources into surface water (Flanders,Belgium).The 6th

international symposium on ecosystem behaviour (BIOGEOMON), Helsinki, Finland, June 29-July 3, 2009. Conference programme and abstracts.

Xue, D., De Baets, B., Botte, J., Van Cleemput, O. and Boeckx, P. Bayesian isotopic mixing model to quantify NO_3^- input sources into surface water (Flanders, Belgium). International workshop "Towards new methods to manage nitrate pollution within the Water Framework Directive" (ISONITRATE), Paris, France, 10-11 December, 2009. Programme and abstracts.

Xue, D., De Baets, B., Botte, J., Vermeulen, J., Van Cleemput, O. and Boeckx, P. Identification and quantification of NO_3^- inputs into surface water in Flanders, Belgium. Benelux Association of Stable Isotope Scientists (BASIS), Brugge, Belgium, 15-16 April, 2010. Programme and abstracts.

Xue, D., De Baets, B., Botte, J., Van Cleemput, O. and Boeckx, P. Bayesian isotopic mixing model to quantify NO_3^- input sources into surface water (Flanders, Belgium). European Geosciences Union General Assembly 2010, Vienna, Austria, 2-7 May, 2010. Programme and abstracts.

Xue, D., De Baets, B., Botte, J., Van Cleemput, O. and Boeckx, P. A dual isotope approach of $\delta^{15}N$- and $\delta^{18}O$-NO_3^- for source identification and quantification of NO_3^- in surface water (Flanders, Belgium). Workshop "Isotopes for improved management of nitrate pollution in aqueous resources" (INPAR), Vienna, Austria, 6 May, 2010. Programme and abstracts.

6. PRESENTATIONS AND POSTERS

International conferences

Benelux Association of Stable Isotope Scientists (BASIS), Arnhem, Netherlands, 3-4 April, 2008. Presentation: Combination of $\delta^{15}N$and $\delta^{18}O$ values of nitrate to identify nitrate sources in Flanders, Belgium.

Benelux Association of Stable Isotope Scientists (BASIS), Brugge, Belgium, 2-3 April, 2009. Presentation: Identification and quantification of NO_3^- inputs into surface water in Flanders, Belgium.

The 6th international symposium on ecosystem behaviour (BIOGEOMON), Helsinki, Finland, June 29-July 3, 2009. Poster: Bayesian isotopic mixing model to quantify NO_3^- input sources into surface water (Flanders, Belgium).

Benelux Association of Stable Isotope Scientists (BASIS), Brugge, Belgium, 15-16 April, 2010. Poster: Identification and quantification of NO_3^- inputs into surface water in Flanders, Belgium.

European Geosciences Union General Assembly 2010, Vienna, Austria, 2-7 May, 2010. Poster: Bayesian isotopic mixing model to quantify NO_3^- input sources into surface water (Flanders, Belgium).

Workshops

International workshop "Towards new methods to manage nitrate pollution within the Water Framework Directive" (ISONITRATE), Paris, France, 10-11 December, 2009. Poster: Bayesian isotopic mixing model to quantify NO_3^- input sources into surface water (Flanders, Belgium).

Workshop "Isotopes for improved management of nitrate pollution in aqueous resources" (INPAR), Vienna, Austria, 6 May, 2010. Presentation: A dual isotope approach of $\delta^{15}N$- and $\delta^{18}O$-NO_3^- for source identification and quantification of NO_3^- in surface water (Flanders, Belgium).